

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 17:44:11 ; Search time 21596 Seconds
(without alignments)
11794.569 Million cell updates/sec

```

Title: US-10-802-228-1
Perfect score: 4481
Sequence: 1 acccccgaaggaaccacccca.....taaaaaaaaaaaaaaaaaaa 4481

```

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s

Database : GenEmbl:*

```
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	4481	100.0	4481	6	AR153580	Sequence	
2	4481	100.0	4481	6	AR401379	Sequence	
3	4481	100.0	4481	6	AR447280	Sequence	
4	4481	100.0	4481	6	AR632098	Sequence	
5	4481	100.0	4481	6	AX392465	Sequence	
6	4481	100.0	4481	8	HSU70323	Human ataxi	
7	4450.6	99.3	4479	6	CQ727533	Sequence	
8	3949	88.1	4163	8	HS2ANS2A	H sapiens m	
9	3949	88.1	4200	6	A62706	Sequence 7	
10	3138.2	70.0	4225	9	AF041472	Mus muscu	
11	2957	66.0	3798	6	AR632099	Sequence	
12	1311.4	29.3	2766	8	AK095017	Homo sapi	
13	1047.8	23.4	2558	5	BC097692	Xenopus l	
14	935.2	20.9	1257	6	AR447281	Sequence	
15	935.2	20.9	1257	6	AR632110	Sequence	
16	935.2	20.9	1257	6	AX392475	Sequence	
17	935.2	20.9	1257	9	MMU70670	Mus musculu	
18	872.2	19.5	107717	8	AC137055	Homo sapi	

C	19	859	19.2	108139	14	AC0040855_3	Continuation (4 of
C	20	731.2	16.3	1080	9	MM278162	278162 M.musculus
C	21	633	14.1	4369	6	CQ851373	CQ851373 Sequence
C	22	633	14.1	4369	8	AK128613	AK128613 Homo sapi
C	23	571.2	12.7	623	6	AR159546	AR159546 Sequence
C	24	497.4	11.1	516	6	AR447279	AR447279 Sequence
C	25	497.4	11.1	516	6	AR632097	AR632097 Sequence
C	26	466	10.4	572	6	AR159558	AR159558 Sequence
C	27	455.8	10.2	459	6	AR401373	AR401373 Sequence
C	28	412.4	9.2	49677	6	AR659776	AR659776 Sequence
C	29	412.4	9.2	155022	8	AC002395	AC002395 Homo sapi
C	30	393.4	8.8	3604	8	AK126309	AK126309 Homo sapi
C	31	389.8	8.7	409	8	AF330029	AF330029 Gorilla g
C	32	365.6	8.2	234319	14	AC160039	AC160039 Bos tauru
C	33	354.2	7.9	355	6	AR159544	AR159544 Sequence
C	34	346.2	7.7	469	10	G07000	G07000 human STS W
C	35	342	7.6	390	8	AF330028	AF330028 Pan trogl
C	36	340.8	7.6	384	8	AF330030	AF330030 prebbytlis
C	37	313.2	7.0	47677	6	AR659627	AR659627 Sequence
C	38	306.6	6.8	67550	14	AC016452	AC016452 Homo sapi
C	39	301.4	6.7	305	6	CQ689654	CQ689654 Sequence
C	40	288.8	6.4	201427	9	AC113302	AC113302 Mus muscu
C	41	287.6	6.4	3764	9	BC054483	BC054483 Mus muscu
C	42	285.8	6.4	1181	9	BC091406	BC091406 Rattus no
C	43	285.8	6.4	228910	14	AC095537	AC095537 Rattus no
C	44	276	6.2	210459	9	AC116104	AC116104 Mus muscu
C	45	269.6	6.0	3758	8	HSAC119794	HSAC119794 Homo sapi

ALIGNMENTS

```

RESULT 1
AR153580
LOCUS AR153580 4481 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 18 from patent US 6235872.
ACCESSION AR153580
VERSION AR153580.1 GI:15121112
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4481)
AUTHORS Predesen,D.E. and Rabizadeh,S.
TITLE Proapoptotic peptides dependence polypeptides and methods of use
JOURNAL Patent: US 6235872-A 18 22-MAY-2001;
FEATURES
source
1. .4481
/organism="unknown"
/mol_type="unassigned DNA"

```

ORIGIN	Query Match	100.0%;	Score 4481;	DB 6;	Length 4481;
	Best Local Similarity	100.0%;	Pred. NO. 0;		
	Matches 4481;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	ACCCCGAGAAAAAGCAACCCAGCGCGCGCGCGCTCCTCAAGTGTCCTCCCGGCCCCCGG	60		
Db	1	ACCCCGAGAAAAAGCAACCCAGCGCGCGCGCGCTCCTCAAGTGTCCTCCCGGCCCCCGG	60		
Qy	61	GCCACCTCAGGTTCTGCTTCGGTCTGACCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG	120		
Db	61	GCCACCTCAGGTTCTGCTTCGGTCTGACCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG	120		
Qy	121	CACCTCCGCTCCACACCGCGGCGCTCAGCGGCGCGCCCTCCGATGCGCTCAGCGGCGCGCA	180		
Db	121	CACCTCCGCTCCACACCGCGGCGCTCAGCGGCGCGCCCTCCGATGCGCTCAGCGGCGCGCA	180		
Qy	181	GCTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCGCAGCCAGGTGGGCC	240		
Db	181	GCTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGCGCAGCCAGGTGGGCC	240		
Qy	241	GGGTGGGCGCTCGTCCAGCGGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG	300		

Db 241 |||||GGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGCGGGCGGGCGGGCGGCGGCC 300
Qy 301 CCGGACCGTATCCCTCCGCGCGCCCTCCCGCGCCCGGCGCCCGCCCTCCCTCCCGG 360
Db 301 CCGGACCGTATCCCTCCGCGCGCCCTCCCGCGCCCGGCGCCCGCCCTCCCTCCCGG 360
Qy 361 CAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAAGCGCAACGGCGGGCGCG 420
Db 361 CAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAAGCGCAACGGCGGGCGCG 420
Qy 421 TTTTGGCGCGGCTCCCGCGGCTCCCTGGTCTCGGCGGCGCTCCCGCGCCCTTGGTCGTC 480
Db 421 TTTTGGCGCGGCTCCCGCGGCTCCCTGGTCTCGGCGGCGCTCCCGCGCCCTTGGTCGTC 480
Qy 481 GTCTCTTCCCTCCG 540
Db 481 GTCTCTTCCCTCCG 540
Qy 541 CTCGGCGCGGCTGGCTCCCGCGCGGTTCCGGCGTCTCTTGGCGCGCGCGCGCGCGCG 600
Db 541 CTCGGCGCGGCTGGCTCCCGCGCGGTTCCGGCGTCTCTTGGCGCGCGCGCGCGCGCG 600
Qy 601 TGTCTCCCGCGCGCTCGAGCGGTGTATGGCGCTCACCATGTCTGCTGAAGCGCCAG 660
Db 601 TGTCTCCCGCGCGCTCGAGCGGTGTATGGCGCTCACCATGTCTGCTGAAGCGCCAG 660
Qy 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Db 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Qy 721 CAGCGCGCGCGCGGCTGCCAATGTCGCAAGCCCGCGCGAGCGCGCTTCTAGCGTCTG 780
Db 721 CAGCGCGCGCGCGGCTGCCAATGTCGCAAGCCCGCGCGAGCGCGCTTCTAGCGTCTG 780
Qy 781 CCCGCGCGCGCGCTTCCGCGCTCTCGTCTCGTCTCTCGTCTCGTCTCGTCTCGTCTCC 840
Db 781 CCCGCGCGCGCGCTTCCGCGCTCTCGTCTCGTCTCTCGTCTCGTCTCGTCTCGTCTCC 840
Qy 841 TCCTCGTGGTCCGCGGACCTCCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 841 TCCTCGTGGTCCGCGGACCTCCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy 901 AGTAACAAGACTGCTCAGTCTAGATTTCTTTGATGAATCTATGGAATATGAGG 960
Db 901 AGTAACAAGACTGCTCAGTCTAGATTTCTTTGATGAATCTATGGAATATGAGG 960
Qy 961 ATGGTTTCATATCTTACATCAGTTGTTGCTGCCAAATGTGAAGTACAGTCAAAATGGA 1020
Db 961 ATGGTTTCATATCTTACATCAGTTGTTGCTGCCAAATGTGAAGTACAGTCAAAATGGA 1020
Qy 1021 GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAGTGTGATTTGGTACTTGTGCGC 1080
Db 1021 GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAGTGTGATTTGGTACTTGTGCGC 1080
Qy 1081 GCATATGAGAAAGTACAGAAATCCAGTTCCGGGCGCGAAACGTGAAGAAATATGAGAGT 1140
Db 1081 GCATATGAGAAAGTACAGAAATCCAGTTCCGGGCGCGAAACGTGAAGAAATATGAGAGT 1140
Qy 1141 ATTTTGTTCAAATGTTTACAGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
Db 1141 ATTTTGTTCAAATGTTTACAGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
Qy 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Qy 1261 GAGAGGACCTTGGAGCCCTGGAGTGCAGGTGAACTCACAGCAATGAGGAACTTGGGCT 1320
Db 1261 GAGAGGACCTTGGAGCCCTGGAGTGCAGGTGAACTCACAGCAATGAGGAACTTGGGCT 1320
Qy 1321 TTGGAAATACGTATCTTAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA 1380

Db 1321 TTGGAAATACGTATCTTAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA 1380
Qy 1381 AATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTTCGTATACAGTGCCTTGA 1440
Db 1381 AATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTTCGTATACAGTGCCTTGA 1440
Qy 1441 AGAGTAACCTCAGAGAAATTTTAAACGGGAGCAAGGCAAAAGGCAAAAGGCAAAAG 1500
Db 1441 AGAGTAACCTCAGAGAAATTTTAAACGGGAGCAAGGCAAAAGGCAAAAGGCAAAAG 1500
Qy 1501 ATTTCAGTCAAGTGCAGTACAAAGCTTCGAGTGGCCCTCGAATAATGATATGAGAGT 1560
Db 1501 ATTTCAGTCAAGTGCAGTACAAAGCTTCGAGTGGCCCTCGAATAATGATATGAGAGT 1560
Qy 1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGACAGCATAAAC 1620
Db 1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGACAGCATAAAC 1620
Qy 1621 ACTAGGGAATAATAATATATTCCTCTGGACAAAGAAATAGAGAAATCATATCTTGGGA 1680
Db 1621 ACTAGGGAATAATAATATATTCCTCTGGACAAAGAAATAGAGAAATCATATCTTGGGA 1680
Qy 1681 AGTGGGAGACAGAAATTCACCGGTATGGCGCAGCTCGATCGGCTCCATGCCATCAAGA 1740
Db 1681 AGTGGGAGACAGAAATTCACCGGTATGGCGCAGCTCGATCGGCTCCATGCCATCAAGA 1740
Qy 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAAATTCGTGTTTCAGACCAAGAGTAGTTAAT 1800
Db 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAAATTCGTGTTTCAGACCAAGAGTAGTTAAT 1800
Qy 1801 GGAGTGTTCCTGGCCATCGCTTGGCCATCTCTCTTCTCTCTCGCCACCTTCTCGCTAC 1860
Db 1801 GGAGTGTTCCTGGCCATCGCTTGGCCATCTCTCTTCTCTCTCGCCACCTTCTCGCTAC 1860
Qy 1861 CAGTCAGTCCCAACTCTCTTCCACCTCGGCGCAGCCCTCACAGCGCGCGCCCTCCAGG 1920
Db 1861 CAGTCAGTCCCAACTCTCTTCCACCTCGGCGCAGCCCTCACAGCGCGCGCCCTCCAGG 1920
Qy 1921 CCCCCCTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCTCATGTTTCTTCCAGCTCT 1980
Db 1921 CCCCCCTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCTCATGTTTCTTCCAGCTCT 1980
Qy 1981 GTCTCTACTATGCTTAAACGATGTCTTCAAGAGGCTTCAAGGATGTCCCAAGGCC 2040
Db 1981 GTCTCTACTATGCTTAAACGATGTCTTCAAGAGGCTTCAAGGATGTCCCAAGGCC 2040
Qy 2041 CAGCGCATCTCTCGAAATCAAGAGTTTCTGCTGGAGGGTTCATATCCAGTGGCTA 2100
Db 2041 CAGCGCATCTCTCGAAATCAAGAGTTTCTGCTGGAGGGTTCATATCCAGTGGCTA 2100
Qy 2101 GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGACCAGT 2160
Db 2101 GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGACCAGT 2160
Qy 2161 CCCTCGGGGGAACGTGGTCTATAGTGGTTCAGTGGGTTCCAGATTTATCCCTTAAACT 2220
Db 2161 CCCTCGGGGGAACGTGGTCTATAGTGGTTCAGTGGGTTCCAGATTTATCCCTTAAACT 2220
Qy 2221 CATAGACCCAGGTTCTCCAGACAGACAGTATTTGGAATACCCCGAGTGGCGAGTCTT 2280
Db 2221 CATAGACCCAGGTTCTCCAGACAGACAGTATTTGGAATACCCCGAGTGGCGAGTCTT 2280
Qy 2281 GCTTCTCCCAAGCTGTTATTTCCAACTGAAGCTGTTCGATGCTTATTCAGTGC 2340
Db 2281 GCTTCTCCCAAGCTGTTATTTCCAACTGAAGCTGTTCGATGCTTATTCAGTGC 2340
Qy 2341 TCTCTAGGCTGTCTAGTCTGTGATCGAAGCAGAGCTGTTACCCCTTCTAGTGGCTAAA 2400
Db 2341 TCTCTAGGCTGTCTAGTCTGTGATCGAAGCAGAGCTGTTACCCCTTCTAGTGGCTAAA 2400
Qy 2401 GATTTCAGGCTTCAAGATCAGAGCAGACTCTCTCTCAGGGAATAAGAAATATTTAAA 2460
Db 2401 GATTTCAGGCTTCAAGATCAGAGCAGACTCTCTCTCAGGGAATAAGAAATATTTAAA 2460

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4481)
AUTHORS Brahmachari,S.K., Choudhry,S., Mukerji,M. and Jain,S.
TITLE Method of detection of allelic variants of SCA2 gene
JOURNAL Patent: US 6623927-A 20 23-SEP-2003;
Council of Scientific and Industrial Research; New Delhi;
INX;
FEATURES
source Location/Qualifiers
1..4481
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 4481; DB 6; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCCCGAGAAAGCAACCCAGGGCGCCGCTCCTCAGGTGTCCCTCCCGGCCCGGG 60
DB 1 ACCCCGAGAAAGCAACCCAGGGCGCCGCTCCTCAGGTGTCCCTCCCGGCCCGGG 60
QY 61 GCCACCTCACGTTCTGCTTCGCTGACCCCTCCGACTTCGGETAAAGAGTCCCTATCCG 120
DB 61 GCCACCTCACGTTCTGCTTCGCTGACCCCTCCGACTTCGGETAAAGAGTCCCTATCCG 120
QY 121 CACCTCCGCTCCACCCGCGCCCTCGGCGCGCCCGCCCTCCGATCGCTCAGCGCGCGCA 180
DB 121 CACCTCCGCTCCACCCGCGCCCTCGGCGCGCCCGCCCTCCGATCGCTCAGCGCGCGCA 180
QY 181 GCTCTCCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGGACGACGAGTGGCC 240
DB 181 GCTCTCCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGGACGACGAGTGGCC 240
QY 241 GGGTGGCGCTGCTCCACGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
DB 241 GGGTGGCGCTGCTCCACGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
QY 301 CCGGACCGGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 CCGGACCGGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CAGAGTCCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGGCGGCGG 420
DB 361 CAGAGTCCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGGCGGCGG 420
QY 421 TTTTGGCGCGGCTCCGCGCGCTCTGCTGCTCGGCGGCGCTCCCGCGCGCGCTTCGTC 480
DB 421 TTTTGGCGCGGCTCCGCGCGCTCTGCTGCTCGGCGGCGCTCCCGCGCGCGCTTCGTC 480
QY 481 GTCTTTCTCCCTCCG 540
DB 481 GTCTTTCTCCCTCCG 540
QY 541 CTCGGCGCGGCTCCCGCGCGGTTCCGGCGTCTCTTGGCGCGCGCGCGCGCGCG 600
DB 541 CTCGGCGCGGCTCCCGCGCGGTTCCGGCGTCTCTTGGCGCGCGCGCGCGCGCG 600
QY 601 TGTCTCCCGCGCGCTGGAGCGGCTGTATGGCGCCCTCACCATGCTGAGCGCCCGAG 660
DB 601 TGTCTCCCGCGCGCTGGAGCGGCTGTATGGCGCCCTCACCATGCTGAGCGCCCGAG 660
QY 661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
DB 661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
QY 721 CAGCGCGCGCGCGCTGCGCAATGTCCGAAGCCCGCGCGCGCGCGCGCGCTTAGCTCG 780
DB 721 CAGCGCGCGCGCGCTGCGCAATGTCCGAAGCCCGCGCGCGCGCGCGCTTAGCTCG 780
QY 781 CCGCGCGCGCGCTTCGCGGCTCTGCTCGGCTCTCTGCTCTCGGCGCGCGCGCTCC 840
DB 781 CCGCGCGCGCGCTTCGCGGCTCTGCTCGGCTCTCTGCTCTCGGCGCGCGCGCTCC 840

QY 841 TCCTCGGTGTGTCGCGCGACCTCCGCGCGCGGAGGCCCGCGCTTGGCAGAGGTGGAAC 900
DB TCCTCGGTGTGTCGCGCGACCTCCGCGCGCGGAGGCCCGCGCTTGGCAGAGGTGGAAC 900
QY 901 AGTACAAAGGACTGCTCAGTCTACGATTCTTTTGGATCGAATCTATGCANAATATGAGG 960
DB AGTACAAAGGACTGCTCAGTCTACGATTCTTTTGGATCGAATCTATGCANAATATGAGG 960
QY 961 ATGGTTCATATACCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAATGGA 1020
DB ATGGTTCATATACCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAATGGA 1020
QY 1021 GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAAGTGTGATTTGGTATTTGATGCC 1080
DB GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAAGTGTGATTTGGTATTTGATGCC 1080
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTCGGGCGCGAAACGTGAAGAAATTAATGGAGGT 1140
DB GCACATGAGAAAGTACAGAAATCCAGTTCGGGCGCGAAACGTGAAGAAATTAATGGAGGT 1140
QY 1141 ATTTTGTTCAAATGTTTTCAGACTTTTGTGTGTACAGTTTAAAGATATGGACTCCAGTTAT 1200
DB ATTTTGTTCAAATGTTTTCAGACTTTTGTGTGTACAGTTTAAAGATATGGACTCCAGTTAT 1200
QY 1201 GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGCGCAACACAAA 1260
DB GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGCGCAACACAAA 1260
QY 1261 GAGAAGACCTGGAGCCCTGGATGCGAGTGAACACACGCAATGAGGACTTTGAGGCT 1320
DB GAGAAGACCTGGAGCCCTGGATGCGAGTGAACACACGCAATGAGGACTTTGAGGCT 1320
QY 1321 TTGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATAATGAAGAA 1380
DB TTGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATAATGAAGAA 1380
QY 1381 AATTATCGGTGAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1440
DB AATTATCGGTGAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1440
QY 1441 AGAGATACTCAGAAAGAAATTTTAAACCGGAGCAAGGCGCAACCCAGTTAGCAGAGAA 1500
DB AGAGATACTCAGAAAGAAATTTTAAACCGGAGCAAGGCGCAACCCAGTTAGCAGAGAA 1500
QY 1501 ATTGAGTCAAAGTGCCTCAGTACAAAGCTCGAGTGGCCCTGGAAAAATGATAGGAGTGAG 1560
DB ATTGAGTCAAAGTGCCTCAGTACAAAGCTCGAGTGGCCCTGGAAAAATGATAGGAGTGAG 1560
QY 1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAAC 1620
DB GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAAC 1620
QY 1621 ACTAGGGAAATAAATATATTTCCCTCGGACAAAGAAATAGAGAGTCAATATCTCGGGA 1680
DB ACTAGGGAAATAAATATATTTCCCTCGGACAAAGAAATAGAGAGTCAATATCTCGGGA 1680
QY 1681 AGTGGGAGACAGAAATTCACCGCGTATGGGCGAGCGCTCGATCGGCTCCATGCCATCAAGA 1740
DB AGTGGGAGACAGAAATTCACCGCGTATGGGCGAGCGCTCGATCGGCTCCATGCCATCAAGA 1740
QY 1741 TCCAATCTTCACTTCAGATTTCGAATTCGAATTCGAGTTCAGAACCAAGAGTAGTTAAT 1800
DB TCCAATCTTCACTTCAGATTTCGAATTCGAATTCGAGTTCAGAACCAAGAGTAGTTAAT 1800
QY 1801 GGAGGTGTTCCTGGCCATCGCCTTGCCTTCCCTCCCTCCGCGCGCGCGCTTCTCGCTAC 1860
DB GGAGGTGTTCCTGGCCATCGCCTTGCCTTCCCTCCCTCCGCGCGCGCGCTTCTCGCTAC 1860
QY 1861 CAGTCAGGTCCCAACTCTCTTCCACCTCGGCGAGCCGCCCTACACGCGCGCGCTCCAGG 1920
DB CAGTCAGGTCCCAACTCTCTTCCACCTCGGCGAGCCGCCCTACACGCGCGCGCTCCAGG 1920

QY 1921 CCCCCCTCGGGGCATCCAGACCCCGCTCTCAGCCCTCTGCTCATGGTTCTCCAGCTCTT 1980
DB 1921 CCCCCCTCGGGGCATCCAGACCCCGCTCTCAGCCCTCTGCTCATGGTTCTCCAGCTCTT 1980
QY 1981 GTCTCTACTATGCTTAAAGCATGTCTTCAGAAAGGCTCTCAAGGATGTCCCAAGGCTC 2040
DB 1981 GTCTCTACTATGCTTAAAGCATGTCTTCAGAAAGGCTCTCAAGGATGTCCCAAGGCTC 2040
QY 2041 CAGCGACATCTCTCGAAATCAGAGATTTCTGTGGGAGGGGTTCANATCCAGTGGCCTA 2100
DB 2041 CAGCGACATCTCTCGAAATCAGAGATTTCTGTGGGAGGGGTTCANATCCAGTGGCCTA 2100
QY 2101 GAATTTGTATCCACAAACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGCCAGT 2160
DB 2101 GAATTTGTATCCACAAACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGCCAGT 2160
QY 2161 CCGCTCGGGGGAAAGTGGTCTCAGTGGGTTCAGTGGGGTTCCAAGATATCCCTAAAGCT 2220
DB 2161 CCGCTCGGGGGAAAGTGGTCTCAGTGGGTTCAGTGGGGTTCCAAGATATCCCTAAAGCT 2220
QY 2221 CATPAGACCCAGGTCTCCAGACAGAAACAGTATTTGGAAATACCCAGTGGGCGCAGTTCTT 2280
DB 2221 CATPAGACCCAGGTCTCCAGACAGAAACAGTATTTGGAAATACCCAGTGGGCGCAGTTCTT 2280
QY 2281 GCTTCTCCAGAGCTGGTATTTATCCAGTGAAGCTGTGGCAGTCTATTCAGCTGCA 2340
DB 2281 GCTTCTCCAGAGCTGGTATTTATCCAGTGAAGCTGTGGCAGTCTATTCAGCTGCA 2340
QY 2341 TCTCTCTACGCTCTAGTCTGATCGATCGAACAGAGCTGTACCCCTCTAGTGAAGCTTAA 2400
DB 2341 TCTCTCTACGCTCTAGTCTGATCGATCGAACAGAGCTGTACCCCTCTAGTGAAGCTTAA 2400
QY 2401 GAFTTCCAGGCTTCAAGATCAGAGCAGAACTCTCTCGAGGGGAATAAGAAATATTTAAA 2460
DB 2401 GAFTTCCAGGCTTCAAGATCAGAGCAGAACTCTCTCGAGGGGAATAAGAAATATTTAAA 2460
QY 2461 CCAATGAACATCACTAGCTTCTCAAGAGCTGAAACCAAGGTATATCACCAGTTGTT 2520
DB 2461 CCAATGAACATCACTAGCTTCTCAAGAGCTGAAACCAAGGTATATCACCAGTTGTT 2520
QY 2521 TCTGACATAGAAACAGATGATGATTTAAAGAAATTTAAAGAAATGATTTAGTTTACAG 2580
DB 2521 TCTGACATAGAAACAGATGATGATTTAAAGAAATTTAAAGAAATGATTTAGTTTACAG 2580
QY 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAATAGAGGGGAGAAAAA 2640
DB 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAATAGAGGGGAGAAAAA 2640
QY 2641 TCAAGAGATTTGATCAAGACAAATTTGAACCAAGTCTTAAGGATTTCTTTTCAATTGAAAT 2700
DB 2641 TCAAGAGATTTGATCAAGACAAATTTGAACCAAGTCTTAAGGATTTCTTTTCAATTGAAAT 2700
QY 2701 AGCAGAGCAACTGTACAGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2760
DB 2701 AGCAGAGCAACTGTACAGTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2760
QY 2761 ATACTTAGTAAACAGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCAGACT 2820
DB 2761 ATACTTAGTAAACAGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCAGACT 2820
QY 2821 TCCAGGCCAGCATGTAAACAGAGAAAGACGATAGAGAAAGAGAAAGACGCGCTGAG 2880
DB 2821 TCCAGGCCAGCATGTAAACAGAGAAAGACGATAGAGAAAGAGAAAGACGCGCTGAG 2880
QY 2881 CAAGTTAGGAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCCAGTCTCTCTCT 2940
DB 2881 CAAGTTAGGAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCCAGTCTCTCTCTCT 2940
QY 2941 CAGCAAAGCTTCTACTACCCCAACTTCACTCTGGGCTCAAGCAGCAACCTTAGCCCCATCT 3000
DB 2941 CAGCAAAGCTTCTACTACCCCAACTTCACTCTGGGCTCAAGCAGCAACCTTAGCCCCATCT 3000
QY 3001 ATGGTGGGTCTATCAACAGCCAACTCCAGTTTATCTCAGGCTGTGTTTGTGACCCAAAT 3060

DB 3001 ATGGTGGGTCTATCAACAGCCAACTCCAGTTTATCTCAGCCTGTTTGTGACCAAAAT 3060
QY 3061 ATGATGTATCCAGTCCCGAGTCCCGAGCGGTGCAACCTTTATACCCAAATACCTATGAGC 3120
DB 3061 ATGATGTATCCAGTCCCGAGTCCCGAGCGGTGCAACCTTTATACCCAAATACCTATGAGC 3120
QY 3121 CCCATGCCAGTGAATCAAGACCAAGACATATAGAGCAGTACCAAAATATGCCCAAGCGG 3180
DB 3121 CCCATGCCAGTGAATCAAGACCAAGACATATAGAGCAGTACCAAAATATGCCCAAGCGG 3180
QY 3181 CAAAGACCAAGCATCATCAGAGTCCATGATGCAACCCAGCGGTTCAGAGCGGCGCCACCGATT 3240
DB 3181 CAAAGACCAAGCATCATCAGAGTCCATGATGCAACCCAGCGGTTCAGAGCGGCGCCACCGATT 3240
QY 3241 GCAGGACCCCAACAGCTTACTCCACGCAATATGTTGCTTACAGTCTCAGCAGTTCCTCA 3300
DB 3241 GCAGGACCCCAACAGCTTACTCCACGCAATATGTTGCTTACAGTCTCAGCAGTTCCTCA 3300
QY 3301 AATCAGCCCTTGTTCAGCATGTGCAATATCAGTCTCAGCATCCTCATCTATAGT 3360
DB 3301 AATCAGCCCTTGTTCAGCATGTGCAATATCAGTCTCAGCATCCTCATCTATAGT 3360
QY 3361 CCTGTAAATACAGGTAATGCTAGAAATGATGGCACCACCAACACAGCCCGCCTGGTTTA 3420
DB 3361 CCTGTAAATACAGGTAATGCTAGAAATGATGGCACCACCAACACAGCCCGCCTGGTTTA 3420
QY 3421 GTATCTTCTTACAGAACTCAGTACGGGCTCATGAGCAGCAGCATGCGATGTATGATGT 3480
DB 3421 GTATCTTCTTACAGAACTCAGTACGGGCTCATGAGCAGCAGCATGCGATGTATGATGT 3480
QY 3481 CCAAAATTTACCATACAAAGGAGACAAGCCCTTCTTCTTACTTTGCCATTTCCACGGGC 3540
DB 3481 CCAAAATTTACCATACAAAGGAGACAAGCCCTTCTTCTTACTTTGCCATTTCCACGGGC 3540
QY 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTTACCTCGACCCACATACTCCACACCT 3600
DB 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTTACCTCGACCCACATACTCCACACCT 3600
QY 3601 CAGCCTTACAGCTACCCCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3660
DB 3601 CAGCCTTACAGCTACCCCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3660
QY 3661 CCGAGTCTCTGTTTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3720
DB 3661 CCGAGTCTCTGTTTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3720
QY 3721 CAGCAGCAGTACAGCAATTTTACCGCGGGCTTGGCCAACTCCACCTCCATGACACCT 3780
DB 3721 CAGCAGCAGTACAGCAATTTTACCGCGGGCTTGGCCAACTCCACCTCCATGACACCT 3780
QY 3781 GCCTCCAAACAGCAGTCCGACAGAAATAGTTTCCAGCAGCAGCAGCAGCAGCAGCAGC 3840
DB 3781 GCCTCCAAACAGCAGTCCGACAGAAATAGTTTCCAGCAGCAGCAGCAGCAGCAGCAGC 3840
QY 3841 ATCCATCTCTTCAAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCACAGTACCT 3900
DB 3841 ATCCATCTCTTCAAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCACAGTACCT 3900
QY 3901 CAGGCTCATGTACAGTACAGCAATGTTTCTCTCATCTCAACTGCGCCATGGCCCAATGATG 3960
DB 3901 CAGGCTCATGTACAGTACAGCAATGTTTCTCTCATCTCAACTGCGCCATGGCCCAATGATG 3960
QY 3961 CTAAATGACACACAGCAGCAGCAGCGGCTCCCGCGGCTCCCTCGCTCAAAAGTGCATACAG 4020
DB 3961 CTAAATGACACACAGCAGCAGCAGCGGCTCCCGCGGCTCCCGCTCGCTCAAAAGTGCATACAG 4020
QY 4021 CCAATTTCCAGTCTCGAACAAGCGCAATTTCCCTCTATATGACGCAACCTTTTTCAGTACAGCC 4080
DB 4021 CCAATTTCCAGTCTCGAACAAGCGCAATTTCCCTCTATATGACGCAACCTTTTTCAGTACAGCC 4080
QY 4081 CACCAACACAGCAGTGTAGGCTCCCTGCGGAGAACCGAAAGGCCAAATTCCTCTCTC 4140

Dn	4081	CACCAACACAGCTGTGAAGCTGCCTGGAGAACCGAAAGGCCAAAATCCCTCCTC	4140	Dn	301	CGGGAACGTAATCCCTCCGCGCCCTCCCGCCCGCGCCCTCCCTCCCTCCCTCCCGG	360
Qy	4141	CCTTCTACTGCTTTACCAACTGGGAAGCACAGAAACTAGAAATTCATTATTTGTTTT	4200	Qy	361	CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGAACGCGGGCGCGCG	420
Dn	4141	CCTTCTACTGCTTTACCAACTGGGAAGCACAGAAACTAGAAATTCATTATTTGTTTT	4200	Dn	361	CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGAACGCGGGCGCGCG	420
Qy	4201	TAAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTCACTTGCG	4260	Qy	421	TTTGGGCCCGCTCCGCGGCTCCTTGGTCTCGGCGGCGCTCCCGCGCCCTTGGTCGT	480
Dn	4201	TAAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTCACTTGCG	4260	Dn	421	TTTGGGCCCGCTCCGCGGCTCCTTGGTCTCGGCGGCGCTCCCGCGCCCTTGGTCGT	480
Qy	4261	TGGAAGATACTTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTCCATAATTCCA	4320	Qy	481	GTCTCTCTCCCTTCGCAGCCCGGCGCCCTCCGCGCGCAACCCGCGCCCTCCCGG	540
Dn	4261	TGGAAGATACTTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTCCATAATTCCA	4320	Dn	481	GTCTCTCTCCCTTCGCAGCCCGGCGCCCTCCGCGCGCAACCCGCGCCCTCCCGG	540
Qy	4321	TATGCTGTTTACAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT	4380	Qy	541	CTCGCGCCCGTGGGTCCCGCCCGCTTCGCGCGCTCTCTTGGCGCGCCCGCTCCCGC	600
Dn	4321	TATGCTGTTTACAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT	4380	Dn	541	CTCGCGCCCGTGGGTCCCGCCCGCTTCGCGCGCTCTCTTGGCGCGCCCGCTCCCGC	600
Qy	4381	ATTTTAAATACCTTGGAAGTCATGACACATCAGCTAGCAAAAGAGTAACAGAGT	4440	Qy	601	TGTCCTCCCGCCGCGTGGAGCGGTATGGGCCCTCACCATGTCTGCTGAAGCCCGAG	660
Dn	4381	ATTTTAAATACCTTGGAAGTCATGACACATCAGCTAGCAAAAGAGTAACAGAGT	4440	Dn	601	TGTCCTCCCGCCGCGTGGAGCGGTATGGGCCCTCACCATGTCTGCTGAAGCCCGAG	660
Qy	4441	GATTCCTGCTGCTATTACTGCTAAAAAATAAAAAAAAAAAAAA	4481	Qy	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720
Dn	4441	GATTCCTGCTGCTATTACTGCTAAAAAATAAAAAAAAAAAAAA	4481	Dn	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720
RESULT 3							
AR447280							
LOCUS AR447280 4481 bp mRNA linear PAT 20-FEB-2004							
DEFINITION Sequence 2 from patent US 6673535.							
ACCESSION AR447280							
VERSION AR447280.1 GI:42675575							
KEYWORDS							
SOURCE Unknown.							
ORGANISM Unknown.							
REFERENCE Unclassified.							
AUTHORS 1 (bases 1 to 4481)							
TITLE Pulat,S.M.							
JOURNAL Methods of detecting spinocerebellar Ataxia-2 nucleic acids							
PATENT: US 6673535-A 2 06-JAN-2004;							
Cedars-Sinai Medical Center, Los Angeles, CA							
FEATURES							
source Location/Qualifiers							
1..4481							
/organism="unknown"							
/mol_type="mRNA"							
ORIGIN							
Query Match 100.0%; Score 4481; DB 6; Length 4481;							
Best Local Similarity 100.0%; Pred. No. 0;							
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;							
Qy	1	ACCCCGAGAAAGCAACCCAGCGCGCGCTCTCAGCTGTCCCTCCCGGCCCGCGG	60	Qy	1081	GCATGAGAAAAGTACAGAAATCCAGTTCCGGGGCCGAAACCGTGAAGAAATATGGAGAGT	1140
Dn	1	ACCCCGAGAAAGCAACCCAGCGCGCGCTCTCAGCTGTCCCTCCCGGCCCGCGG	60	Dn	1081	GCATGAGAAAAGTACAGAAATCCAGTTCCGGGGCCGAAACCGTGAAGAAATATGGAGAGT	1140
Qy	61	GCCACCTCAGCTTCTGCTTCCGCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG	120	Qy	1141	ATTTGTTTCAAATGTTTACAGCTTTGTTGTGTAGAGTTTAAAGATATGGAATCCAGTTAT	1200
Dn	61	GCCACCTCAGCTTCTGCTTCCGCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG	120	Dn	1141	ATTTGTTTCAAATGTTTACAGCTTTGTTGTGTGTAGAGTTTAAAGATATGGAATCCAGTTAT	1200
Qy	121	CACCTCCGCTCCACCCGCGGCTCGGGCGGCCCGCTCCGATGGCTCAGCGCGCGCA	180	Qy	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGGCAACACAAA	1260
Dn	121	CACCTCCGCTCCACCCGCGGCTCGGGCGGCCCGCTCCGATGGCTCAGCGCGCGCA	180	Dn	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGGCAACACAAA	1260
Qy	181	GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCCAGAGTGGGCC	240	Qy	1261	GAGAAGACCTGGAGCCCTGGGATGCGAGTGAATCACAGCAATGAGGAATCTGAGGCT	1320
Dn	181	GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCCAGAGTGGGCC	240	Dn	1261	GAGAAGACCTGGAGCCCTGGGATGCGAGTGAATCACAGCAATGAGGAATCTGAGGCT	1320
Qy	241	GGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGGCGGGTGGCGCGGCC	300	Qy	1321	TTGGAAAAATGACGATCTTAATGGATGGGATCCCAATGATATGTTTCGATATTAATGAAGAA	1380
Dn	241	GGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGGCGGGTGGCGCGGCC	300	Dn	1321	TTGGAAAAATGACGATCTTAATGGATGGGATCCCAATGATATGTTTCGATATTAATGAAGAA	1380
Qy	301	CGGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGCGGCCCGGCCCTCCCTCCCGG	360	Qy	1381	AAATATGGGTAGTGTCTACGATGATAGAGTTTATCTTCGTATACAGTCCCTTACAA	1440

Db 1381 AATTATGGTGTAGTGTCTACGTAATGATAGCAGTTTATCTTCGTATACAGTGCCCTTAGAA 1440
Qy 1441 AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAGAA 1500
Db 1441 AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAGAA 1500
Qy 1501 ATTGAGTCAAGTCCCGCTACAAAGCTCGAGTGGCCCTGGAAATGATAGAGAGTGG 1560
Db 1501 ATTGAGTCAAGTCCCGCTACAAAGCTCGAGTGGCCCTGGAAATGATAGAGAGTGG 1560
Qy 1561 GAAAGAAAATACACAGCAGTTCAGAGAAATCCAGTGAACGTGAGGGGCACAGCATAAAC 1620
Db 1561 GAAAGAAAATACACAGCAGTTCAGAGAAATCCAGTGAACGTGAGGGGCACAGCATAAAC 1620
Qy 1621 ACTAGGGAAAATPAAATATATCTCTCTGGCAAAAGAAATAGAGAAAGTCATATCTCTGGGGA 1680
Db 1621 ACTAGGGAAAATPAAATATATCTCTCTGGCAAAAGAAATAGAGAAAGTCATATCTCTGGGGA 1680
Qy 1681 AGTGGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCGGGCTCCATGCCATCAAGA 1740
Db 1681 AGTGGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCGGGCTCCATGCCATCAAGA 1740
Qy 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAATCTCTGGTTCAAGCCAAAGAGTAGTTAAT 1800
Db 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAATCTCTGGTTCAAGCCAAAGAGTAGTTAAT 1800
Qy 1801 GGAGTGTCTCCCTGGCCATCGGCTTGGCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
Db 1801 GGAGTGTCTCCCTGGCCATCGGCTTGGCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
Qy 1861 CAGTCAGTCCCAACTCTCTTCCACTCGGGCAGCCACCCCTACAGGGCGGCCCTCCAGG 1920
Db 1861 CAGTCAGTCCCAACTCTCTTCCACTCGGGCAGCCACCCCTACAGGGCGGCCCTCCAGG 1920
Qy 1921 CCCCCCTCGCGCATCAGAGCCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGTCTCT 1980
Db 1921 CCCCCCTCGCGCATCAGAGCCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGTCTCT 1980
Qy 1981 GTCTCTACTATGCTTAAACGCATGTCTTCAGAGGGCCCTCCAAGGATGTCCCAAGGCC 2040
Db 1981 GTCTCTACTATGCTTAAACGCATGTCTTCAGAGGGCCCTCCAAGGATGTCCCAAGGCC 2040
Qy 2041 CAGCGACATCTCGAAATCAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCCTA 2100
Db 2041 CAGCGACATCTCGAAATCAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCCTA 2100
Qy 2101 GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGACCACT 2160
Db 2101 GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGACCACT 2160
Qy 2161 CCCTCGGGGGGAACGTGGTCACTCAGTGGTCACTGAGTGGGTTCCAAGATTTCCCTTAAACT 2220
Db 2161 CCCTCGGGGGGAACGTGGTCACTCAGTGGTCACTGAGTGGGTTCCAAGATTTCCCTTAAACT 2220
Qy 2221 CATAGACCGAGTCTCCAGACAGACAGTATTTGGAAATACCCAGTGGGCGCAGTTCTT 2280
Db 2221 CATAGACCGAGTCTCCAGACAGACAGTATTTGGAAATACCCAGTGGGCGCAGTTCTT 2280
Qy 2281 GCTTCTCCCCAGCTGCTATTTTCAACTGAAGCTGTGGCATGCTATTCAGCTGCA 2340
Db 2281 GCTTCTCCCCAGCTGCTATTTTCAACTGAAGCTGTGGCATGCTATTCAGCTGCA 2340
Qy 2341 TCTCTCTACGCTGTAGTCTCGATCGAACAGAGCTGTTTACCCCTTCTAGTGGGCTAAA 2400
Db 2341 TCTCTCTACGCTGTAGTCTCGATCGAACAGAGCTGTTTACCCCTTCTAGTGGGCTAAA 2400
Qy 2401 GATTCAGGCTTCAAGATCAGAGGCGAATCTCTCTCGAGGGGAATTAAGAAATATATAA 2460
Db 2401 GATTCAGGCTTCAAGATCAGAGGCGAATCTCTCTCGAGGGGAATTAAGAAATATATAA 2460
Qy 2461 CCAATGAACATCACTAGCTTCTCAAGCTGAAACCAAGGTATATCACCAGTTGT 2520
Db 2461 CCAATGAACATCACTAGCTTCTCAAGCTGAAACCAAGGTATATCACCAGTTGT 2520

Qy 2521 TCTGAACATAGAAAAACAGATTTGATGATTTAAAGAAATTTAAAGAAATGATTTTAGTTACAG 2580
Db 2521 TCTGAACATAGAAAAACAGATTTGATGATTTAAAGAAATTTAAAGAAATGATTTTAGTTACAG 2580
Qy 2581 CCAAGTTCTACTCTCTGAACTATGATCAACTACTATAAACAATAAATAGAGAGGAGAAAA 2640
Db 2581 CCAAGTTCTACTCTCTGAACTATGATCAACTACTATAAACAATAAATAGAGAGGAGAAAA 2640
Qy 2641 TCAAGAGATTTGATCAAAAGACAAAATTTGAACCAAGTGTAAAGGATTTCTTCAATTGAAAAAT 2700
Db 2641 TCAAGAGATTTGATCAAAAGACAAAATTTGAACCAAGTGTAAAGGATTTCTTCAATTGAAAAAT 2700
Qy 2701 AGCAGCAGCAACTGTACCACTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA 2760
Db 2701 AGCAGCAGCAACTGTGTACCACTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA 2760
Qy 2761 ATACTTTAGTAACACGGAGCAACAGAGGGGACCTGAGTCACTTTCCCAAGGGGTTCCAGACT 2820
Db 2761 ATACTTTAGTAACACGGAGCAACAGAGGGGACCTGAGTCACTTTCCCAAGGGGTTCCAGACT 2820
Qy 2821 TCCAGCCAGCATGTAAACAGAGAGAAAGACGATAGGAAGAGAGAAAGACGACCTGAG 2880
Db 2821 TCCAGCCAGCATGTAAACAGAGAGAAAGACGATAGGAAGAGAGAAAGACGACCTGAG 2880
Qy 2881 CAAAGTTAGGAAATCAACATTGAAATCCCAATGAAATGAAAGSAGTTCAACCCAGTTCTCTCT 2940
Db 2881 CAAAGTTAGGAAATCAACATTGAAATCCCAATGAAATGAAAGSAGTTCAACCCAGTTCTCTCT 2940
Qy 2941 CAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCTCAAGCAACACCTTAGCCCATCT 3000
Db 2941 CAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCTCAAGCAACACCTTAGCCCATCT 3000
Qy 3001 ATGGTGGTCTATCAACAGCCAACTCCAGTTTATCTCAGCCCTGTTGTTTGGCACAAT 3060
Db 3001 ATGGTGGTCTATCAACAGCCAACTCCAGTTTATCTCAGCCCTGTTGTTTGGCACAAT 3060
Qy 3061 ATGATGTATCCAGTCCCGAGTGGCCAGCGTGCACCTTTATACCAATACCTATGAGC 3120
Db 3061 ATGATGTATCCAGTCCCGAGTGGCCAGCGTGCACCTTTATACCAATACCTATGAGC 3120
Qy 3121 CCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAATATGCCCCAACAGCG 3180
Db 3121 CCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAATATGCCCCAACAGCG 3180
Qy 3181 CAAAGACCAGCATCATCAGATGTCATGATGCAACCCAGCGTCAAGCGGGGCCACCGATT 3240
Db 3181 CAAAGACCAGCATCATCAGATGTCATGATGCAACCCAGCGTCAAGCGGGGCCACCGATT 3240
Qy 3241 GCAGCCACCCACAGCTTACTCCAGGCATATGTTGCTACAGTCTCCTCAGCAGTTCCCA 3300
Db 3241 GCAGCCACCCACAGCTTACTCCAGGCATATGTTGCTACAGTCTCCTCAGCAGTTCCCA 3300
Qy 3301 AATCAGCCCTGTTTTCAGATGTCACATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
Db 3301 AATCAGCCCTGTTTTCAGATGTCACATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
Qy 3361 CCTGTAATACAGGGTAATGCTAGAAATGAGCAACCAACACAGCCCGCAGCTTGTTTA 3420
Db 3361 CCTGTAATACAGGGTAATGCTAGAAATGAGCAACCAACACAGCCCGCAGCTTGTTTA 3420
Qy 3421 GTATCTTCTTACGAACTCAGTACGGGCTCATGAGCAGACGATGCGATGTATGATGT 3480
Db 3421 GTATCTTCTTACGAACTCAGTACGGGCTCATGAGCAGACGATGCGATGTATGATGT 3480
Qy 3481 CCCAAATTAACATCAACAGGAGACAGCCCTTCTTCTACTTTTGGCCATTTCCACGGGC 3540
Db 3481 CCCAAATTAACATCAACAGGAGACAGCCCTTCTTCTACTTTTGGCCATTTCCACGGGC 3540
Qy 3541 TCCCTTGCTCAGCAGTATGCGCAACCTTAACGCTACCTCGACCCACATACTCCACACCT 3600
Db 3541 TCCCTTGCTCAGCAGTATGCGCAACCTTAACGCTACCTCGACCCACATACTCCACACCT 3600

QY 3601 CAGCCTTCAGCTACCCCTGAGCAGCAGCAAGCCCAACATGCTGGAAGTCATCCTGCA 3660
Db |||||
QY 3601 CAGCCTTCAGCTACCCCTGAGCAGCAGCAAGCCCAACATGCTGGAAGTCATCCTGCA 3660
Db |||||
QY 3661 CCAGCTCCTGTTCCAGCACCATCAGCACCAGGCGCGCCAGGCTCTCATCTGCGCCAGTCCA 3720
Db |||||
QY 3661 CCAGCTCCTGTTCCAGCACCATCAGCACCAGGCGCGCCAGGCTCTCATCTGCGCCAGTCCA 3720
Db |||||
QY 3721 CAGCAGCAGTACGCAATTTACACGCGGGCTTTGCGCCAACTCCACCTCCATGACACCT 3780
Db |||||
QY 3721 CAGCAGCAGTACGCAATTTACACGCGGGCTTTGCGCCAACTCCACCTCCATGACACCT 3780
Db |||||
QY 3781 GCCTCCAAACGCACTGCGCAGAGATAGTTTCCAGCAGCAGCAACAGACTGCTTTTACG 3840
Db |||||
QY 3781 GCCTCCAAACGCACTGCGCAGAGATAGTTTCCAGCAGCAGCAACAGACTGCTTTTACG 3840
Db |||||
QY 3841 ATCCATCTCTTCAGCTTCAGCAGCGGCTATACCAACCCACATGCGCCACCTACT 3900
Db |||||
QY 3841 ATCCATCTCTTCAGCTTCAGCAGCGGCTATACCAACCCACATGCGCCACCTACT 3900
Db |||||
QY 3901 CAGGCTCATGTACAGTACGGAATGTTCTTCTCATCCAACTGCCCATGCGCCAAATGATG 3960
Db |||||
QY 3901 CAGGCTCATGTACAGTACGGAATGTTCTTCTCATCCAACTGCCCATGCGCCAAATGATG 3960
Db |||||
QY 3961 CTAAATGAGCAGCAGCAGCCCGGCTCCAGGCGCGCTCGCTCAAAGTGCACTACAG 4020
Db |||||
QY 3961 CTAAATGAGCAGCAGCAGCCCGGCTCCAGGCGCGCTCGCTCAAAGTGCACTACAG 4020
Db |||||
QY 4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTCATATGACGACCCCTTCAGTACAAGCC 4080
Db |||||
QY 4021 CCCATTCCAGTCTCGACAAACAGCGCATTTCCCTCATATGACGACCCCTTCAGTACAAGCC 4080
Db |||||
QY 4081 CACCAACCAACAGCAGTGTAAAGCTGCCCTGAGGAACCGAAAGGCCAAATTCCTCTCTC 4140
Db |||||
QY 4081 CACCAACCAACAGCAGTGTAAAGCTGCCCTGAGGAACCGAAAGGCCAAATTCCTCTCTC 4140
Db |||||
QY 4141 CTTCTACTGCTTCTACCACTGGAGCAGCAGAACTAGAAATTCATTATTTTGTGTTT 4200
Db |||||
QY 4141 CTTCTACTGCTTCTACCACTGGAGCAGCAGAACTAGAAATTCATTATTTTGTGTTT 4200
Db |||||
QY 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAAACAGTTTCACTTGCGAG 4260
Db |||||
QY 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAAACAGTTTCACTTGCGAG 4260
Db |||||
QY 4261 TGGAGATACCTTGGACCGAGTAGAGGCAATTTAGGAACTTTGGGGCTATTTCCATTAATCCA 4320
Db |||||
QY 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGGCGAACTGGAAGTTATTT 4380
Db |||||
QY 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGGCGAACTGGAAGTTATTT 4380
Db |||||
QY 4381 ATTTTATTAATACCTTCGAAAGTCAATGAACATCAGCTAGCAAGAAAGTAAACAAGAGT 4440
Db |||||
QY 4381 ATTTTATTAATACCTTCGAAAGTCAATGAACATCAGCTAGCAAGAAAGTAAACAAGAGT 4440
Db |||||
QY 4441 GATTCTGCTGCTATTAATGCTTAAATAAAAAAAAAAAAAA 4481
Db |||||
QY 4441 GATTCTGCTGCTATTAATGCTTAAATAAAAAAAAAAAAAA 4481
Db |||||

RESULT 4
LOCUS AR632098
DEFINITION Sequence 2 from patent US 684431.
ACCESSION AR632098
VERSION AR632098.1
KEYWORDS GI:59773824
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4481)
AUTHORS Pulst,S.M.

TITLE Nucleic acid encoding spinocerebellar ataxia-2 and products related thereto
JOURNAL Patent: US 684431-A 2 18-JAN-2005;
Cedars-Sinai Medical Center; Los Angeles, CA
FEATURES
source 1. 4481
/organism="unknown"
/mol_type="mRNA"
ORIGIN
Query Match 100.0%; Score 4481; DB 6; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCCCCCGAGAAAGCAACCCAGCGCGCCCGCTCCTCACGTGTCTCCTCCCGGCCCCGGG 60
Db |||||
QY 1 ACCCCCCGAGAAAGCAACCCAGCGCGCCCGCTCCTCACGTGTCTCCTCCCGGCCCCGGG 60
Db |||||
QY 61 GCGACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120
Db |||||
QY 61 GCGACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120
Db |||||
QY 121 CACCTCCGCTCCACCCGCGCTCGGCGCGCCCGCTCCGATGCGCTCAGCGCGCGCA 180
Db |||||
QY 121 CACCTCCGCTCCACCCGCGCTCGGCGCGCCCGCTCCGATGCGCTCAGCGCGCGCA 180
Db |||||
QY 181 GCTCCTCGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGGCC 240
Db |||||
QY 181 GCTCCTCGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGGCC 240
Db |||||
QY 241 GGTGGCGCTCGCTCAGCGCGCGCGCGGAGCGCGCGGGCGGGCGGTGGCGCGGC 300
Db |||||
QY 241 GGTGGCGCTCGCTCAGCGCGCGCGCGGAGCGCGCGGGCGGGCGGTGGCGCGGC 300
Db |||||
QY 301 CCGGACCGTATCCCTCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db |||||
QY 301 CCGGACCGTATCCCTCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db |||||
QY 361 CAGAGCTCGCTCCCTCGCGCTCAGACTGTTTGGTAGCAACCGGCGCGCGCGCG 420
Db |||||
QY 361 CAGAGCTCGCTCCCTCGCGCTCAGACTGTTTGGTAGCAACCGGCGCGCGCGCG 420
Db |||||
QY 421 TTTCCGCGCGCTCCCGCGCTCTTGGTCTCGCGCGCGCTCCCGCGCGCTTGGTCTGC 480
Db |||||
QY 421 TTTCCGCGCGCTCCCGCGCTCTTGGTCTCGCGCGCGCTCCCGCGCGCTTGGTCTGC 480
Db |||||
QY 481 GTCTTCTCCCTCG 540
Db |||||
QY 481 GTCTTCTCCCTCG 540
Db |||||
QY 541 CTCGCGCGCGTGGCTCCCGCGCTTCGCGCGCTCTCTTGGCGCGCGCGCGCTCCCGCG 600
Db |||||
QY 541 CTCGCGCGCGTGGCTCCCGCGCTTCGCGCGCTCTCTTGGCGCGCGCGCGCTCCCGCG 600
Db |||||
QY 601 TGTCCCGCGCGCGTGGAGCGGTGTATGGCGCTTCAACATGTGCTGAAGCGCGCG 660
Db |||||
QY 601 TGTCCCGCGCGCGTGGAGCGGTGTATGGCGCTTCAACATGTGCTGAAGCGCGCG 660
Db |||||
QY 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAG 720
Db |||||
QY 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAG 720
Db |||||
QY 721 CAGCGCGCGCGCGGTGCGCAATGTCCGCAAGCGCGCGCGCGCGCGCGCTTCTAGCGTGC 780
Db |||||
QY 721 CAGCGCGCGCGCGGTGCGCAATGTCCGCAAGCGCGCGCGCGCGCGCTTCTAGCGTGC 780
Db |||||
QY 781 CCCGCGCGCGCGCTTCCGCGTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCC 840
Db |||||
QY 781 CCCGCGCGCGCGCTTCCGCGTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCC 840
Db |||||
QY 841 TCCTCGGTGGTGGCGGCGCTCCGCGCGCGGAGGCGCGCGCTGGGCGAGAGTTCGAAC 900
Db |||||
QY 841 TCCTCGGTGGTGGCGGCGCTCCGCGCGCGGAGGCGCGCGCTGGGCGAGAGTTCGAAC 900
Db |||||

QY 901 AGTAAAGAGAGCTGCTCAGTCTAGATTTCTTTTGATGGAATCTATGCCAAATATGAGG 960
DB 901 AGTAAAGAGAGCTGCTCAGTCTAGTCTACGATTTCTTTTGATGGAATCTATGCCAAATATGAGG 960
QY 961 ATGGTTTATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAAGAAATGGA 1020
DB 961 ATGGTTTATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAAGAAATGGA 1020
QY 1021 GGTATATATGAAGGAGTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1080
DB 1021 GGTATATATGAAGGAGTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1080
QY 1081 GCACATGAGAAAGATCAGAAATCCAGTTCGGGGCCGAAACGTCGAGAAATATGAGAGT 1140
DB 1081 GCACATGAGAAAGATCAGAAATCCAGTTCGGGGCCGAAACGTCGAGAAATATGAGAGT 1140
QY 1141 ATTTTGTTCAAATGTTTCAGACTTTGTTGTGTACAGTTTAAAGATATGAGACTCCAGTTAT 1200
DB 1141 ATTTTGTTCAAATGTTTCAGACTTTGTTGTGTACAGTTTAAAGATATGAGACTCCAGTTAT 1200
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCGAAACACAAA 1260
DB 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCGAAACACAAA 1260
QY 1261 GAGAAAGACTGGAGCCCTGGGATGAGGTGAACTCACGCCAAATGAGAACTTTGAGGCT 1320
DB 1261 GAGAAAGACTGGAGCCCTGGGATGAGGTGAACTCACGCCAAATGAGAACTTTGAGGCT 1320
QY 1321 TTGGAAATGACGTATCTAATGATGGATCCCAATGATATGTTTCGATATATGAGAA 1380
DB 1321 TTGGAAATGACGTATCTAATGATGGATCCCAATGATATGTTTCGATATATGAGAA 1380
QY 1381 AATTATGTTGATGTTCTAGTATGATAGCAGTTTATCTTCGTATACAGTGGCTTTAGAA 1440
DB 1381 AATTATGTTGATGTTCTAGTATGATAGCAGTTTATCTTCGTATACAGTGGCTTTAGAA 1440
QY 1441 AGAGATAACTCAGAGAAATTTTAAACCGGGAAGCAAGGCAACACAGTTAGCAGAGAA 1500
DB 1441 AGAGATAACTCAGAGAAATTTTAAACCGGGAAGCAAGGCAACACAGTTAGCAGAGAA 1500
QY 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATCATATAGGAGTGAG 1560
DB 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATCATATAGGAGTGAG 1560
QY 1561 GAAGAAAAATACACAGCAGTTTACAGAAATTTCCAGTGAACGTCGAGGGGCAACAGATAAAC 1620
DB 1561 GAAGAAAAATACACAGCAGTTTACAGAAATTTCCAGTGAACGTCGAGGGGCAACAGATAAAC 1620
QY 1621 ACTAGGGAATAAATATATTTCTCTCGGACAAAGAAATAGAGAGTCTATCTCGGGA 1680
DB 1621 ACTAGGGAATAAATATATTTCTCTCGGACAAAGAAATAGAGAGTCTATCTCGGGA 1680
QY 1681 AGTGGAGACAGAAATTCACCGGTATGGGCGAGCCTGGATCGGGCTCCATGCCATCAAGA 1740
DB 1681 AGTGGAGACAGAAATTCACCGGTATGGGCGAGCCTGGATCGGGCTCCATGCCATCAAGA 1740
QY 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAAATTCCTGGTTTCAGACCAAGAGTAGTTAAT 1800
DB 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAAATTCCTGGTTTCAGACCAAGAGTAGTTAAT 1800
QY 1801 GGAAGTGTTCCTCGGCGCATCGCCTTGCCTATCTCTCTCTCGCCCACTTCTCGCTAC 1860
DB 1801 GGAAGTGTTCCTCGGCGCATCGCCTTGCCTATCTCTCTCTCGCCCACTTCTCGCTAC 1860
QY 1861 CAGTCAGGTCCCAACTCTCTTTCACCTTCGGGCGACCAACCCCTACACGGCGCCCTCAGG 1920
DB 1861 CAGTCAGGTCCCAACTCTCTTTCACCTTCGGGCGACCAACCCCTACACGGCGCCCTCAGG 1920
QY 1921 CCCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGTCTCT 1980
DB 1921 CCCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGTCTCT 1980

QY 1981 GTCTCTACTATGCTTAAACGATGTCTTTCAGAGGGCCTTCAAGGATGTCCCAAGGCC 2040
DB 1981 GTCTCTACTATGCTTAAACGATGTCTTTCAGAGGGCCTTCAAGGATGTCCCAAGGCC 2040
QY 2041 CAGCCACATCTCTCGAAATCAGAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100
DB 2041 CAGCCACATCTCTCGAAATCAGAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100
QY 2101 GAATTTGTATCCCAACACCCAGTCAGAGCTACTCTCCAGTAGCAGGACCACT 2160
DB 2101 GAATTTGTATCCCAACACCCAGTCAGAGCTACTCTCCAGTAGCAGGACCACT 2160
QY 2161 CCTCGGGGGGAACTGTGTCTCAGTGTCTCAGTGGGTTTCAAGATTTATCCCTTAAACT 2220
DB 2161 CCTCGGGGGGAACTGTGTCTCAGTGTCTCAGTGGGTTTCAAGATTTATCCCTTAAACT 2220
QY 2221 CATAGACCCAGTCTCCAGACAGAACAGTATTTGAAATATACCCAGTGGGCGAGTTCT 2280
DB 2221 CATAGACCCAGTCTCCAGACAGAACAGTATTTGAAATATACCCAGTGGGCGAGTTCT 2280
QY 2281 GCTTCTCCCAAGCTGGTATTTATTTCAACTGAACTGTGGCATGCTATTTCCAGTGCA 2340
DB 2281 GCTTCTCCCAAGCTGGTATTTATTTCAACTGAACTGTGGCATGCTATTTCCAGTGCA 2340
QY 2341 TCTCTAGCCCTGTAGTCTGATCGAACAGAGCTGTTTACCCCTTCTAGTGAGGCTAA 2400
DB 2341 TCTCTAGCCCTGTAGTCTGATCGAACAGAGCTGTTTACCCCTTCTAGTGAGGCTAA 2400
QY 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAGAAATATTTAA 2460
DB 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAGAAATATTTAA 2460
QY 2461 CCCAATGAAACATCACCTAGTCTTCTCAAAAGCTGAAACAAAGGTATATCACAGTTG 2520
DB 2461 CCCAATGAAACATCACCTAGTCTTCTCAAAAGCTGAAACAAAGGTATATCACAGTTG 2520
QY 2521 TCTGAACATAGAAACAGATTTGATGATTTTAAAGAAATTTTAAAGATGATTTAGTTACAG 2580
DB 2521 TCTGAACATAGAAACAGATTTGATGATTTTAAAGAAATTTTAAAGATGATTTAGTTACAG 2580
QY 2581 CCAAGTTCTACTTCTGAATCTATGATCAACTACTTAAACAAAGGTATAGAGGGAGAAAA 2640
DB 2581 CCAAGTTCTACTTCTGAATCTATGATCAACTACTTAAACAAAGGTATAGAGGGAGAAAA 2640
QY 2641 TCAAGAGATTTGATCAAGACAAAAATTGAACCAAGTGTCTTCAATTTGAAAT 2700
DB 2641 TCAAGAGATTTGATCAAGACAAAAATTGAACCAAGTGTCTTCAATTTGAAAT 2700
QY 2701 AGCAGCAGCACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCGAGCATTTCCCTTCA 2760
DB 2701 AGCAGCAGCACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCGAGCATTTCCCTTCA 2760
QY 2761 ATACTTAGTAAACAGGAGCACAAGGGGACCTGAGGTCACTTTCCCAAGGGTTTCAGACT 2820
DB 2761 ATACTTAGTAAACAGGAGCACAAGGGGACCTGAGGTCACTTTCCCAAGGGTTTCAGACT 2820
QY 2821 TCCAGCCAGCATGTAAACAGAGAAAGACGATAGAGAAAGAGAAAGACGAGCTGAG 2880
DB 2821 TCCAGCCAGCATGTAAACAGAGAAAGACGATAGAGAAAGAGAAAGACGAGCTGAG 2880
QY 2881 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTCT 2940
DB 2881 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTCT 2940
QY 2941 CAGCCAAAGCTTCTACTACCCCACTTCACTCGGCTTCAAGCAAACTAGCCCATCT 3000
DB 2941 CAGCCAAAGCTTCTACTACCCCACTTCACTCGGCTTCAAGCAAACTAGCCCATCT 3000
QY 3001 ATGGTGGGTCTCAACAGCCCACTCCAGTTTATCTCAGCCTGTTTGTGTCACCAAT 3060
DB 3001 ATGGTGGGTCTCAACAGCCCACTCCAGTTTATCTCAGCCTGTTTGTGTCACCAAT 3060
QY 3061 ATGATGATTCAGTCCAGTGAGCCCGAGGCTGCAACCTTTTATACCAATATCTATGACG 3120

Db	3061		ATGATGATCCAGTCCAGTGGAGCCAGGCGTGCAACTTTTATACCAATACCTATGAAG	3120
Qy	3121		CCCATGCCAGTGAATCAAGCCAGACATATAGAGCAGTACAAATATGCCCAACAGCGG	3180
Db	3121		CCCATGCCAGTGAATCAAGCCAGACATATAGAGCAGTACAAATATGCCCAACAGCGG	3180
Qy	3181		CAAGACAGCATCATCAGAGTGCCATGATGACCCAGCGTCAGCAGGGGCCCAACGATTT	3240
Db	3181		CAAGACAGCATCATCAGAGTGCCATGATGACCCAGCGTCAGCAGGGGCCCAACGATTT	3240
Qy	3241		GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCCTACAGTCTCAGCAGTTCCCA	3300
Db	3241		GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCCTACAGTCTCAGCAGTTCCCA	3300
Qy	3301		AATCAGGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGT	3360
Db	3301		AATCAGGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGT	3360
Qy	3361		CTGTATAACAGGGTAATGCTAGAAATGATGGCAACCAACACACGCCCCAGCTGTGTTTA	3420
Db	3361		CTGTATAACAGGGTAATGCTAGAAATGATGGCAACCAACACACGCCCCAGCTGTGTTTA	3420
Qy	3421		GTATCTTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGATGCGATGTATGATGT	3480
Db	3421		GTATCTTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGATGCGATGTATGATGT	3480
Qy	3481		CCCAAAATACCATAACAAGGAGACAAAGCCCTTCTTTCTACTTTGCCATTTCCAGGGC	3540
Db	3481		CCCAAAATACCATAACAAGGAGACAAAGCCCTTCTTTCTACTTTGCCATTTCCAGGGC	3540
Qy	3541		TCCCTTGCTCAGCAGTATGGCGCACCTTAAGCTACCCCTGCACCCACATACTCCACACCT	3600
Db	3541		TCCCTTGCTCAGCAGTATGGCGCACCTTAAGCTACCCCTGCACCCACATACTCCACACCT	3600
Qy	3601		CAGCCTTCAGTTACCCCACTGGACAGCAGCAAAGCCAAATGCTGGAAGTCACTCTGCA	3660
Db	3601		CAGCCTTCAGTTACCCCACTGGACAGCAGCAAAGCCAAATGCTGGAAGTCACTCTGCA	3660
Qy	3661		CCAGTCCCTGTTACAGCACCATCAGCACCGCGCGCCAGGCTCTCCATCTGGCGCAGTCCA	3720
Db	3661		CCAGTCCCTGTTACAGCACCATCAGCACCGCGCGCCAGGCTCTCCATCTGGCGCAGTCCA	3720
Qy	3721		CAGCAGCAGTCAGCATTATACACGCGGGGCTTGGCGCAACTCCACCCCTCCATGACACCT	3780
Db	3721		CAGCAGCAGTCAGCATTATACACGCGGGGCTTGGCGCAACTCCACCCCTCCATGACACCT	3780
Qy	3781		GCCTCCAAACGAGTCCGACAGATAGTTTCCAGCAGGACAAAGAGCTGCTTTTACG	3840
Db	3781		GCCTCCAAACGAGTCCGACAGATAGTTTCCAGCAGGACAAAGAGCTGCTTTTACG	3840
Qy	3841		ATCCATCCTTCTCAGTTACGCGGCTATACCAACCCACCCACATGGCCCGCAGTACCT	3900
Db	3841		ATCCATCCTTCTCAGTTACGCGGCTATACCAACCCACCCACATGGCCCGCAGTACCT	3900
Qy	3901		CAGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAACTGCCCATGGCCAAATGATG	3960
Db	3901		CAGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAACTGCCCATGGCCAAATGATG	3960
Qy	3961		CTAATGACGACACGCAACCGCGCGTCCCGAGCCGCTTCGCTCAAAGTGCACTACAG	4020
Db	3961		CTAATGACGACACGCAACCGCGCGTCCCGAGCCGCTTCGCTCAAAGTGCACTACAG	4020
Qy	4021		CCCATTCAGTCTCGACAAAGCGCATTTCCCTCATATGACGACCCCTCAGTACAGCC	4080
Db	4021		CCCATTCAGTCTCGACAAAGCGCATTTCCCTCATATGACGACCCCTCAGTACAGCC	4080
Qy	4081		CACCACCAACAGCAGTTGTAAAGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCCTC	4140
Db	4081		CACCACCAACAGCAGTTGTAAAGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCCTC	4140
Qy	4141		CTTCTACTGCTTCTACAACTGGAAGCAGACAGAAACTAGAAATTCATTATTTGTTTT	4200
Db	4141		CTTCTACTGCTTCTACAACTGGAAGCAGACAGAAACTAGAAATTCATTATTTGTTTT	4200

Db	4141	CCTTCTACTGCTCTTCAACAACTGGAAGACAGAAAACTAGAAATTTCAATTTATTTTGTGTTT	4200
Qy	4201	TAAATATATATATCTTGATTTCTTTGTAACTCAATAGGAATGCTAACAGTTCACTTGCAG	4260
Db	4201	TAAATATATATGTTGATTTCTTTGTAACTCAATAGGAATGCTAACAGTTCACTTGCAG	4260
Qy	4261	TGGAAGATACTTGGACCGAGTAGAGGCAATTAGGAACCTTGGGGGCTATTTCCTAAATTTCCA	4320
Db	4261	TGGAAGATACTTGGACCGAGTAGAGGCAATTAGGAACCTTGGGGGCTATTTCCTAAATTTCCA	4320
Qy	4321	TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTTGGAGGTATTATTT	4380
Db	4321	TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTTGGAGGTATTATTT	4380
Qy	4381	ATTTTTTAATAACCCCTTGAAGTCATGAACACATCAGCTAGCAAAAAAGAGTAACAAGAGT	4440
Db	4381	ATTTTTTAATAACCCCTTGAAGTCATGAACACATCAGCTAGCAAAAAAGAGTAACAAGAGT	4440
Qy	4441	GATCTTGCTGCTATTACTGCTTAAAAAATAAAAAAAAAAAAAA 4481	
Db	4441	GATCTTGCTGCTATTACTGCTTAAAAAATAAAAAAAAAAAAAA 4481	
RESULT 5	AX392465		
LOCUS	AX392465	4481 bp	DNA linear PAT 23-MAR-2002
DEFINITION	Sequence 1 from Patent WO0216417.		
ACCESSION	AX392465		
VERSION	AX392465.1	GI:19700764	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Eumarchontoglires; Primates; Catarrhini;		
	Hominoidea; Homo.		
REFERENCE	1		
AUTHORS	Pulst, S.M. and Huynh, D.P.		
TITLE	Transgenic animal model for neurodegenerative disease and uses thereof		
JOURNAL	Patent: WO 0216417-A 1 28-FEB-2002;		
FEATURES	CEDARS-SINAI MEDICAL CENTER (US)		
source	Location/Qualifiers		
	1. .4481		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	163. .4101		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD29012.1"		
	/db_xref="GI:19700765"		
	/translation="MPSAAAAPRSPAVATESRRFAAARWPGWRS LQRPARRSGRGGGG AAPGPPPSAAPPFGPPPPSPKQSSADCFGNSNGGAGFRGCRLLGLUGGPR PPFVILLPLASGAPAAPATPRASPLGARSPRSVSLARPAPGPRCPRAEGLUGGPR MSLKTPQQQQQQQQQQQQQQPPPAANVRKPGSGLLASPAAPASPSSSV SSSSATAPSSVVAATSGGGRGLGRGNSKGLPOSTISFDGIYANMRVHILTSVVG SKCEVOVKNGLIYGVFKTYSKCDLVLDAAHEKSTESSSGPKRERIMESILFKCSDF VVQFQDMSSYAKRDADFTDSAISAKVNGEHEKLEDPWDAEGLTANELEALENDVS NGPDMDFRNEENTGVSTYDSSLSYVTPLEDRNSBEFLKREANOLAEIESS AQYKVALENDDRSEERKTYAVRNSREGHSINTRENKVIIPQQRNREVISWGS RQNSPMGQFGSGMSRSTSTSDFNPSGSDPRTVNGVGVPWSPSPSPSPSPSP QSGNSLPLPRAATPTTPRPSRPSRPPSPAHGSPAPVSTMPKRMSESGPMPSP QAGHPRNHRVSAGRSISGLLEFVSHNPSEAA TTPVARTSPSGCTWSSVVGVPRI. SPKTHPRSPRQNSIGNTSGPVLASPOAGIIPTEAVANPIPAASTPASPNRAVT PSEAKDRLQDRQNSPAGNKNIKNPETSFSFKAENKGISPVVSEHRKQIIDLK KFNDFLPQSPSTSESDQLLNKRNKEKSRDLTKKIEPSAKDSFLENSSNCTSGSS KNSPSGISLPSLITLTHKRGVPTSGQVOTSPACKQEKDKEEKDAAEQVRKSTLN PKAKENPNSFSQKPTTTPSPRQAOQSPSMVGHQPTPVYTOPVCFAPNMYPVP VSPGVOLYPIPMTPMNVCAKTYRAVPMNPOQDQHQHSAAMHPASAAGPPIAATP PAYSTQYAEVSPQFNQPLQVHPYQSHFVSPVYQGNARMWAPPTHAQPGLV SSATQYGAETHAMTAKPLPYNKETSFSFYAISTGLAQYAHNPATLHPHTPH QPSATQQQSQHQSHQSPVTHPSHHQAAQLHALSPQOOSAIYHAGLAATPPSM TPASNTQSQNSFPAAQPTVTHPSHHQVAYTNPPHNAHPVQAHVQSGMVPSPHTAH		
CDS			

ORIGIN									
APMMLMTTQPPGGPQALAQSAQLQIPVSTTAHPFYMTHPSPVQAHQOOL"									
Query Match 100.0%; Score 4481; DB 6; Length 4481;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ACCCCGAGAAAGCAACCCAGCGCGCCCGCTCTCACTGTCCTCTCCCGCGCCCGCGG	60						
Db	1	ACCCCGAGAAAGCAACCCAGCGCGCCCGCTCTCACTGTCCTCTCCCGCGCCCGCGG	60						
Qy	61	GCCACCTCAGCTTCGCTTCGGTCTGACCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG	120						
Db	61	GCCACCTCAGCTTCGCTTCGGTCTGACCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG	120						
Qy	121	CACCTTCGCTCCACCCCGCGCTCCGGCGCGCCCGCTCCGATCGCTCAGCGCCGCA	180						
Db	121	CACCTTCGCTCCACCCCGCGCTCCGGCGCGCCCGCTCCGATCGCTCAGCGCCGCA	180						
Qy	181	GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCCGACGACGAGTGGGCC	240						
Db	181	GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCCGACGACGAGTGGGCC	240						
Qy	241	GGGTGGCGTCCCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	300						
Db	241	GGGTGGCGTCCCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	300						
Qy	301	CGGAGACGTATCCCTCCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG	360						
Db	301	CGGAGACGTATCCCTCCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG	360						
Qy	361	CAGAGCTCCCTCCCTCCCGCTCAGACTGTTTGGTGAAGCAAGCGCGCGCGCGCG	420						
Db	361	CAGAGCTCCCTCCCTCCCGCTCAGACTGTTTGGTGAAGCAAGCGCGCGCGCGCG	420						
Qy	421	TTTCCGCGCGCTCCCGCGCGCTTCTGGTCTCGCGCGCGCTCCCGCGCGCTTCTGCTC	480						
Db	421	TTTCCGCGCGCTCCCGCGCGCTTCTGGTCTCGCGCGCGCTCCCGCGCGCTTCTGCTC	480						
Qy	481	GTCTTTCTCCCTCCCGCGCGCGCTCCCGCGCGCGCTCCCGCGCGCTCCCGCGCG	540						
Db	481	GTCTTTCTCCCTCCCGCGCGCGCTCCCGCGCGCGCTCCCGCGCGCTCCCGCGCG	540						
Qy	541	CTCGCGCGCGTGGTCCCGCGCGCTCCCGCGCTTCTGGCGCGCGCGCGCGCGCG	600						
Db	541	CTCGCGCGCGTGGTCCCGCGCGCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCG	600						
Qy	601	TGTCCTCCCGCGCTTCGAGCGCGTGTATGGCCCTCACCATGTCGCTGAAGCCCGAG	660						
Db	601	TGTCCTCCCGCGCTTCGAGCGCGTGTATGGCCCTCACCATGTCGCTGAAGCCCGAG	660						
Qy	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG	720						
Db	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG	720						
Qy	721	CAGCGCGCGCGCTCCCAATGTCCGCAAGCCCGCGCGCGCGCGCGCGCTTCTAGCGTG	780						
Db	721	CAGCGCGCGCGCTCCCAATGTCCGCAAGCCCGCGCGCGCGCGCGCGCTTCTAGCGTG	780						
Qy	781	CCCGCGCGCGCTTCGCGCTTCGCTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGT	840						
Db	781	CCCGCGCGCGCTTCGCGCTTCGCTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGT	840						
Qy	841	TCCTCGGTGGTCCGCGGACCTCCGCGCGCGGAGCCCGCGCTGGGAGAGGTGGAAC	900						
Db	841	TCCTCGGTGGTCCGCGGACCTCCGCGCGCGGAGCCCGCGCTGGGAGAGGTGGAAC	900						
Qy	901	AGTAAACAAAGGACTGCTCAGTACGATTTCTTTTGTAGTGAATCTATGCAAAATATGAGG	960						
Db	901	AGTAAACAAAGGACTGCTCAGTACGATTTCTTTTGTAGTGAATCTATGCAAAATATGAGG	960						
Qy	961	ATGGTTTCATATCTACATCAGTTGTTGGTCCCAATGTGAAGTACAGTGAAGATGGA	1020						
Db	961	ATGGTTTCATATCTACATCAGTTGTTGGTCCCAATGTGAAGTACAGTGAAGATGGA	1020						
Qy	1021	GGTATATATGAAGGAGTTTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC	1080						
Db	1021	GGTATATATGAAGGAGTTTTTAAACCTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC	1080						
Qy	1081	GCACATGAGAAAGTACAGAAATCCAGTTCGGGGCGGAACCTGAAGAAATATCGAGAGT	1140						
Db	1081	GCACATGAGAAAGTACAGAAATCCAGTTCGGGGCGGAACCTGAAGAAATATCGAGAGT	1140						
Qy	1141	ATTTTGTTCAAATGTTTACAGCTTTGTTGTGTACAGTTTAAAGATATGAGACTCCAGTTAT	1200						
Db	1141	ATTTTGTTCAAATGTTTACAGCTTTGTTGTGTGTACAGTTTAAAGATATGAGACTCCAGTTAT	1200						
Qy	1201	GCAAAAAGAGATGCTTTTACGACTCTGCTATCAGTCTAAAGTGAATGCGGCAACACAAA	1260						
Db	1201	GCAAAAAGAGATGCTTTTACGACTCTGCTATCAGTCTAAAGTGAATGCGGCAACACAAA	1260						
Qy	1261	GAGAGGACCTGGAGCCCTGGGATGCGAGTCAACTCACAGCAATGAGGAACTTCGAGCT	1320						
Db	1261	GAGAGGACCTGGAGCCCTGGGATGCGAGTCAACTCACAGCAATGAGGAACTTCGAGCT	1320						
Qy	1321	TTGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATAAGAGAA	1380						
Db	1321	TTGAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATAAGAGAA	1380						
Qy	1381	AATTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1440						
Db	1381	AATTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	1440						
Qy	1441	AGAGATAACTCAGAGAAATTTTAAACGGGCAACGAGGGCAACACAGTTAGCACAAGAA	1500						
Db	1441	AGAGATAACTCAGAGAAATTTTAAACGGGCAACGAGGGCAACACAGTTAGCACAAGAA	1500						
Qy	1501	ATTGAGTCAAGTGGCCAGTCAAAAGCTCGAGTGGCCCTGGAAAATGATAGAGAGTGA	1560						
Db	1501	ATTGAGTCAAGTGGCCAGTCAAAAGCTCGAGTGGCCCTGGAAAATGATAGAGAGTGA	1560						
Qy	1561	GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGACAGCATAAAC	1620						
Db	1561	GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGACAGCATAAAC	1620						
Qy	1621	ACTAGGAAAAATTAATATATCTCTCGACAAAGAAATAGAGATCATATCTCTGGGGA	1680						
Db	1621	ACTAGGAAAAATTAATATATCTCTCGACAAAGAAATAGAGATCATATCTCTGGGGA	1680						
Qy	1681	AGTGGGAGACAGAAATTCACCGCTATGGCCAGCCTGGATCGGGCTCCATGCCATCAAGA	1740						
Db	1681	AGTGGGAGACAGAAATTCACCGCTATGGCCAGCCTGGATCGGGCTCCATGCCATCAAGA	1740						
Qy	1741	TCCACTTCTCACACTTCAGATTTCAACCGAAATTCGTTTCAGACAAAGAGTAGTTAAT	1800						
Db	1741	TCCACTTCTCACACTTCAGATTTCAACCGAAATTCGTTTCAGACAAAGAGTAGTTAAT	1800						
Qy	1801	GGAGTGTTCCTGGCCATCGCTTGGCCATCTCTCTCGCCACCTCTCGCCCTCTCGCTAC	1860						
Db	1801	GGAGTGTTCCTGGCCATCGCTTGGCCATCTCTCTCGCCACCTCTCGCCCTCTCGCTAC	1860						
Qy	1861	CAGTCAGTCCCAACTCTCTCCACTCGGGCAGCAGCCTTACACGGCGCCCTCCAGG	1920						
Db	1861	CAGTCAGTCCCAACTCTCTCCACTCGGGCAGCAGCCTTACACGGCGCCCTCCAGG	1920						
Qy	1921	CCCCCTCGCGCCATCCAGACCCCGCTCAACCCTCTGCTCATGGTCTCCAGTCTCT	1980						
Db	1921	CCCCCTCGCGCCATCCAGACCCCGCTCAACCCTCTGCTCATGGTCTCCAGTCTCT	1980						
Qy	1981	GTCTCTACTATGCTTAAAGCATGTCTTCAGAGGGCTTCCAGGATGTCCTCCAAAGGCC	2040						
Db	1981	GTCTCTACTATGCTTAAAGCATGTCTTCAGAGGGCTTCCAGGATGTCCTCCAAAGGCC	2040						
Qy	2041	CAGGCACATCTTCGAAAATCAGAGAGTTTCTGCTGGAGGGGTTCATATATCCAGTGGCCTA	2100						
Db	2041	CAGGCACATCTTCGAAAATCAGAGAGTTTCTGCTGGAGGGGTTCATATATCCAGTGGCCTA	2100						

Qy	2101	GAATTTGTATCCCAAAACCCACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGACCACT	2160
Ds	2101	GAATTTGTATCCCAAAACCCACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGACCACT	2160
Qy	2161	CCCTCGGGGGGAAGTGTGTATCATAGTGTGAGTGGGTTCAGAGTATATCCCTTAAACT	2220
Ds	2161	CCCTCGGGGGGAAGTGTGTATCATAGTGTGAGTGGGTTCAGAGTATATCCCTTAAACT	2220
Qy	2221	CATAGACCCAGGCTCTCCAGACAGAACTAGTATGGAAATACCCCAAGTGGGCCAGTTCCT	2280
Ds	2221	CATAGACCCAGGCTCTCCAGACAGAACTAGTATGGAAATACCCCAAGTGGGCCAGTTCCT	2280
Qy	2281	GCTTCTCCCAAGCTGTATTAATCCAACTGAAGCTGTGGCATGCTTATCCAGCTGCA	2340
Ds	2281	GCTTCTCCCAAGCTGTATTAATCCAACTGAAGCTGTGGCATGCTTATCCAGCTGCA	2340
Qy	2341	TCTCTAGCCCTGTAGTCTGTCATCGAAGCAGCTGTATCCCTTCTAGTGAGGCTAAA	2400
Ds	2341	TCTCTAGCCCTGTAGTCTGTCATCGAAGCAGCTGTATCCCTTCTAGTGAGGCTAAA	2400
Qy	2401	GATTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCGCAGGGAATAAGAAAAATATTA	2460
Ds	2401	GATTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCGCAGGGAATAAGAAAAATATTA	2460
Qy	2461	CCCAATGAAACATCACTAGCTTCTCAAAAGCTGAAAAAATAAGGATATATCAACG	2520
Ds	2461	CCCAATGAAACATCACTAGCTTCTCAAAAGCTGAAAAAATAAGGATATATCAACG	2520
Qy	2521	TCTGAACATAGAAACAGATGATGATTTAAGAAATTTAAGAAATTTAAGGTTACAG	2580
Ds	2521	TCTGAACATAGAAACAGATGATGATTTAAGAAATTTAAGAAATTTAAGGTTACAG	2580
Qy	2581	CCAAGTTCTACTTCTGAACTATGATCAACTACTTAAACAAAAATAGAGGGGAAAA	2640
Ds	2581	CCAAGTTCTACTTCTGAACTATGATCAACTACTTAAACAAAAATAGAGGGGAAAA	2640
Qy	2641	TCAAGAGATTTGATCAAAAGAAATTTGAAACAAAGTCTAAGGATCTTTTCAATGAA	2700
Ds	2641	TCAAGAGATTTGATCAAAAGAAATTTGAAACAAAGTCTAAGGATCTTTTCAATGAA	2700
Qy	2701	AGCAGCAGCACTGTACAGTGGCAGCAGCAGCGGATAGCCCAAGATTTCCCTTCA	2760
Ds	2701	AGCAGCAGCACTGTACAGTGGCAGCAGCAGCGGATAGCCCAAGATTTCCCTTCA	2760
Qy	2761	ATACTTAGTAAACAGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCA	2820
Ds	2761	ATACTTAGTAAACAGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCA	2820
Qy	2821	TCCAGCCAGCATGTAAACAGAGAAAGACGATAGGGAAGAGAAAGACGAGCTGAG	2880
Ds	2821	TCCAGCCAGCATGTAAACAGAGAAAGACGATAGGGAAGAGAAAGACGAGCTGAG	2880
Qy	2881	CAAGTTAGGAAATCAACATTCGATCCCAATGCAAGGATTTCAACCCAGTTCCTCT	2940
Ds	2881	CAAGTTAGGAAATCAACATTCGATCCCAATGCAAGGATTTCAACCCAGTTCCTCT	2940
Qy	2941	CAGCCTAAGCCTTCTACTACCCCACTTCACTGGGCTCAAGCACAACCTAGCCCAT	3000
Ds	2941	CAGCCTAAGCCTTCTACTACCCCACTTCACTGGGCTCAAGCACAACCTAGCCCAT	3000
Qy	3001	ATGGTGGGTATCAACAGCCTTCAAGTATATCAAGCTGTGTTGTTGTTGTTGTTGTT	3060
Ds	3001	ATGGTGGGTATCAACAGCCTTCAAGTATATCAAGCTGTGTTGTTGTTGTTGTTGTT	3060
Qy	3061	ATGATGTATCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT	3120
Ds	3061	ATGATGTATCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT	3120
Qy	3121	CCCATGCCAGTGAATCAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAG	3180
Ds	3121	CCCATGCCAGTGAATCAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAG	3180

Qy	3181	CAAGACAGCATCATCAGAGTGCCATGATGCAACCCAGGGTCAGCAGGGGCCACCGATT	3240
Ds	3181	CAAGACAGCATCATCAGAGTGCCATGATGCAACCCAGGGTCAGCAGGGGCCACCGATT	3240
Qy	3241	GCAGCCACCCACAGCTTACTCCAGCAATATGCTGCTTACGTACGTCTCAGCAGTCCCA	3300
Ds	3241	GCAGCCACCCACAGCTTACTCCAGCAATATGCTGCTTACGTACGTCTCAGCAGTCCCA	3300
Qy	3301	AATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCTCATGTCTATAGT	3360
Ds	3301	AATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCTCATGTCTATAGT	3360
Qy	3361	CCTGTAATACAGGTAATGTAGAAATGATGSCAACACACAGCCAGCTGCTGTTTA	3420
Ds	3361	CCTGTAATACAGGTAATGTAGAAATGATGSCAACACACAGCCAGCTGCTGTTTA	3420
Qy	3421	GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCATGCGATGTATGATGT	3480
Ds	3421	GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCATGCGATGTATGATGT	3480
Qy	3481	CCCAAAATTAACATACAAAGGAGCAAGCCCTTCTTCTTACTTTGCCATTTCCACGGGC	3540
Ds	3481	CCCAAAATTAACATACAAAGGAGCAAGCCCTTCTTCTTACTTTGCCATTTCCACGGGC	3540
Qy	3541	TCCCTTGTCTCAGCATGTGCGCAACCTTAAACGCTACCTTGCAACCAATCTCAACCCCT	3600
Ds	3541	TCCCTTGTCTCAGCATGTGCGCAACCTTAAACGCTACCTTGCAACCAATCTCAACCCCT	3600
Qy	3601	CAGCCTTCAAGTACCCCACTGGACAGCAAGCAAGCAATGCTGGAAGTCATCTGCA	3660
Ds	3601	CAGCCTTCAAGTACCCCACTGGACAGCAAGCAAGCAATGCTGGAAGTCATCTGCA	3660
Qy	3661	CCCAAGTCTTCTCAGCACCACAGCAGCCCGCCAGGCTCTCCATCTGCGCCAGTCCA	3720
Ds	3661	CCCAAGTCTTCTCAGCACCACAGCAGCCCGCCAGGCTCTCCATCTGCGCCAGTCCA	3720
Qy	3721	CAGCAGCATGTCAGCATTTTACCAGCGGGGTTGCGCAACTCCACCTTCCATGACACT	3780
Ds	3721	CAGCAGCATGTCAGCATTTTACCAGCGGGGTTGCGCAACTCCACCTTCCATGACACT	3780
Qy	3781	GCCTCCAAACAGCAGTCCGCAAGATAGTTTCCAGCAGCAACAGACTGCTTTTACG	3840
Ds	3781	GCCTCCAAACAGCAGTCCGCAAGATAGTTTCCAGCAGCAACAGACTGCTTTTACG	3840
Qy	3841	ATCCATCTTCTCAGTTCAGCGGGGTTGCGCAACTCCACCTTCCATGACACT	3900
Ds	3841	ATCCATCTTCTCAGTTCAGCGGGGTTGCGCAACTCCACCTTCCATGACACT	3900
Qy	3901	CAGGCTCATGTACAGTCAAGGATGCTTCTCATCCAACTGCGCCATGATGATG	3960
Ds	3901	CAGGCTCATGTACAGTCAAGGATGCTTCTCATCCAACTGCGCCATGATGATG	3960
Qy	3961	CTAATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	4020
Ds	3961	CTAATGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	4020
Qy	4021	CCCAATTCAGTCTCGAACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	4080
Ds	4021	CCCAATTCAGTCTCGAACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	4080
Qy	4081	CACCAACAGCAGTGTAGGCTGCGCTGGAGGAAACCGAAAGGCGCAAAATTCCTCTC	4140
Ds	4081	CACCAACAGCAGTGTAGGCTGCGCTGGAGGAAACCGAAAGGCGCAAAATTCCTCTC	4140
Qy	4141	CTTCTACTCTCTACCACTGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	4200
Ds	4141	CTTCTACTCTCTACCACTGGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	4200
Qy	4201	TAAATATATATGTTGATTTCTTGTAACTCCAAATAGGAAATGCTTAACAGTTTCA	4260
Ds	4201	TAAATATATATGTTGATTTCTTGTAACTCCAAATAGGAAATGCTTAACAGTTTCA	4260
Qy	4261	TGGAAGATACCTTGGACCGAGTAGAGGCACTTGGAGCACTTGGGGGCTATTCCAT	4320

```
|||||
4261 TGAAGATACCTTGGACCGAGTAGAGCATTTAGGAACCTTGGGGCTATTTCCTATAATTC 4320
QY 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380
Db 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380
QY 4381 ATTTTATAAACCCTTGAAAGTCAATGAAACATCATGCTAGCAAAAGAAAGTAAACAAGAGT 4440
Db 4381 ATTTTATAAACCCTTGAAAGTCAATGAAACATCATGCTAGCAAAAGAAAGTAAACAAGAGT 4440
QY 4441 GATCTTCTGCTATTACTGCTAAATAAAAAAAAAAAAAAAAAAAAA 4481
Db 4441 GATCTTCTGCTATTACTGCTAAATAAAAAAAAAAAAAAAAAAAAA 4481

RESULT 6
HSU70323 4481 bp mRNA linear PRI 20-NOV-1996
LOCUS Human ataxin-2 (SCA2) mRNA, complete cds.
DEFINITION U70323
ACCESSION U70323
VERSION 1 GI:1679683
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 4481)
AUTHORS Pulst, S.-M., Nechiporuk, A., Nechiporuk, T., Gispert, S., Chen, X.-N.,
Lopes-Cendes, I., Pearlman, S., Starkman, S., Orozco-biaz, G.,
Lunkes, A., DeJong, P., Rouleau, G.A., Auburger, G., Korenberg, J.R.,
Figueroa, C. and Sahba, S.
TITLE Moderate expansion of a normally biallelic trinucleotide repeat in
spinocerebellar ataxia type 2
JOURNAL Nat. Genet. 14 (3), 269-276 (1996)
PUBMED 8896555
REFERENCE 2 (bases 1 to 4481)
AUTHORS Pulst, S.-M.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1996) Medicine, Cedars-Sinai, 8700 Beverly Blvd.,
Los Angeles, CA 90048, USA
FEATURES
source
1..4481
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q24.1"
1..4481
/genes="SCA2"
163..4101
/genes="SCA2"
/standard_name="spinocerebellar ataxia type 2"
/codon_start=1
/product="ataxin-2"
/protein_id="AAB19200.1"
/db_xref="GI:1679684"
/translation="MRSAAAPRS PAVATESRRFAARWPGBLORPARSRGRGGG
AARGPPSAAPP PPGPPSRQSSPPSASDCFGSNGNGGAPRPSRLILGLGPPR
PFVVLLPLASGAPPAAPTRASPLGARASPLRSGVSLAKPFCPRPACPEVYGLT
MSLKPQQQQQQQQQQQQQQQQQQPPPAANVRKPGSGLLASPAAPSPSSSV
SSSASAPSVAAATGGGRPLGRGRNSKGLPQSTISPDGIYANRMVHLITSVVG
SKCEVQVNGGIYGVGFKYSPKCDLVLDAAHEKSTESSGPKREIMESILPKCDF
VVQPKDMSYAKRDAFTDSAISAKVNGRHKDLPEWDAGELTANEELEALENDVS
NGDPNDMFRYENGVWSTYDSSLSSTVPLERDNESEFLKREARANQLAEEISS
AQYKRALVNDRESEEEKYAVORNSREHSGINTRENKYPPOQRREVISWGG
QNSPRMGQPSGMPSTSHSTDFNPNSSQDQVVGVPWPSPCPSPSPSPSPRY
QSGPNSLPPTAAPTPTPPSPRPPSPSHGSPAPVSTMPKRMSEGPSPRSP
KAQHPNHRVAGRGISISGLEFVSHNPPEAATPPVATSPSGGTWSSVSVSPRL
SKTPHPRSPRONSIGNTPSGVPLASPAQGIIPTEAVAMPIPAASPTPASPNRAVT
PSSEAKDSRLQDRQNSPAGNKENIKPNETSPSFKENKGISPVVSEHRKQIDDLK
FKNDRFLQPSSTSEMDQLLNKREGEKSKDLIDKTEPSAKDSFIENSNSNCTSGSS
```

```
KNPSPISTSLNTEHKRGPEVTSQVOTSSPACKOBKDDKEEKDAAEVOVRKSTLN
PNAKEPNRPSFSQPKPSTTTPSRPAQPSBMVCHQOPTPVYTPVCFAPNMVYVP
VSPGVQPLIPMTMPVNVQAKTYTRAVPNMQRODDHQSNMHPASAGPPIAATP
PAYSTQVAYVSPQFPNQLVQHVHVSQHPHYSPVIOGNARMMAPPTTHAOPGLVS
SSATQYGAHEQTHAMVACPLKPYNKETSPSYFAISTGLSQAOQAHNATLHPHTPH
QPSATPTGQOQSHGGSHAPSPVQHQQHQAALHLASPOQSAIYHAGLAPTPSPM
TPASNTQSPQSPFAAQOTVFTIHPSHVQPAYTNPMMAHVPOAHVQSGMVPSPHTAH
APMMLMTQPPGGPQAALQAQALQIPVSTTAHFPYMTHTPSVQAHHQOOL"
ORIGIN
Query Match 100.0%; Score 4481; DB 8; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTGTCTCCCGCCCCCGG 60
Db 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTGTCTCCCGCCCCCGG 60
QY 61 GCCACCTCACGTTCTGCTTCCGTCCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
Db 61 GCCACCTCACGTTCTGCTTCCGTCCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
QY 121 CACCTCCGCTCCCAACCCGCGCGCTCGCGCGCGCGCTCCGATCGGTCAGCGCGCGCA 180
Db 121 CACCTCCGCTCCCAACCCGCGCGCTCGCGCGCGCGCTCCGATCGGTCAGCGCGCGCA 180
QY 181 GCTCTTCGAGTTCCTCGCGGTGCGCACCGAGTCTCGCGGTTCGCCGACGCCAGGTGGCCC 240
Db 181 GCTCTTCGAGTTCCTCGCGGTGCGCACCGAGTCTCGCGGTTCGCCGACGCCAGGTGGCCC 240
QY 241 GGGTGGCGCTCCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 GGGTGGCGCTCCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 CCGGACACCTATACCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 CCGGACACCTATACCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CAGAGTCCGCTCCTTCGCGCTCAGACTGTTTGTGTAGCAAGCGCGCGCGCGCGCG 420
Db 361 CAGAGTCCGCTCCTTCGCGCTCAGACTGTTTGTGTAGCAAGCGCGCGCGCGCGCG 420
QY 421 TTTTGGCGCGCGCTCCCGCGCGCTCTTGTGTCTCGCGCGCGCTCCCGCGCGCTTCGTCGTC 480
Db 421 TTTTGGCGCGCGCTCCCGCGCGCTCTTGTGTCTCGCGCGCGCTCCCGCGCGCTTCGTCGTC 480
QY 481 GTCTTCTCTCCCTCGCCAGCGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCTCCCGCG 540
Db 481 GTCTTCTCTCCCTCGCCAGCGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCTCCCGCG 540
QY 541 CTGCGCGCGCTGCGTCCCGCGCGGTTCGCGGTCTCTTGGCGCGCGCGCGCTCCCGCGCG 600
Db 541 CTGCGCGCGCTGCGTCCCGCGCGGTTCGCGGTCTCTTGGCGCGCGCGCGCTCCCGCGCG 600
QY 601 TGTCCCGCGCGCGCTGCGAGCGGTGTATGGCGCGCTTCCACCATGTCGCTGAAGCGCGCG 660
Db 601 TGTCCCGCGCGCGCTGCGAGCGGTGTATGGCGCGCTTCCACCATGTCGCTGAAGCGCGCG 660
QY 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Db 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
QY 721 CAGCCCGCGCGCGCTGCGCAATGTTCGCAAGCGCGCGCGCGCGCGCGCTTCTTAGCGTCG 780
Db 721 CAGCCCGCGCGCGCTGCGCAATGTTCGCAAGCGCGCGCGCGCGCGCGCTTCTTAGCGTCG 780
QY 781 CCGCGCGCGCGCTTTCGCGCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCTCG 840
Db 781 CCGCGCGCGCGCTTTCGCGCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCTCG 840
QY 841 TCCTTCGTTGTCGCGCGCGACTCCGCGCGCGGAGCGCGCGCTTGGCGAGAGTTCGAAAC 900
Db 841 TCCTTCGTTGTCGCGCGCGACTCCGCGCGCGGAGCGCGCGCTTGGCGAGAGTTCGAAAC 900
```


QY	901	AGTAA	CAAAAGGACTGCCTCAGTCTACGATTTCTTTTGATGGAAATCTATGCAAAATATGAGG	960	
Db	901	AGTAA	CAAAAGGACTGCCTCAGTCTACGATTTCTTTTGATGGAAATCTATGCAAAATATGAGG	960	
QY	961	ATGGTTCATATACCTACATCAGTTGCTGGCTCCAAATGTGAAGTCAAGTGAATAATGGA	1020		
Db	961	ATGGTTCATATACCTACATCAGTTGCTGGCTCCAAATGTGAAGTCAAGTGAATAATGGA	1020		
QY	1021	GGTATATATGAAGGAGTTTAAAACTTACAGTCCGAAGTGTGATTTGGTACTCTGATGCC	1080		
Db	1021	GGTATATATGAAGGAGTTTAAAACTTACAGTCCGAAGTGTGATTTGGTACTCTGATGCC	1080		
QY	1081	GCACATGAGAAAGTACAGAAATCCAGTTCCGGGCCGGAACCGTGAAGAAATATGAGAGGT	1140		
Db	1081	GCACATGAGAAAGTACAGAAATCCAGTTCCGGGCCGGAACCGTGAAGAAATATGAGAGGT	1140		
QY	1141	ATTTTGTTCMAATGTTACAGCTTTGCTGGTACAGTTTAAAGATATGACTCCAGTTAT	1200		
Db	1141	ATTTTGTTCMAATGTTACAGCTTTGCTGGTACAGTTTAAAGATATGACTCCAGTTAT	1200		
QY	1201	GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTAAAGTGAATGGCGAACAACAA	1260		
Db	1201	GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTAAAGTGAATGGCGAACAACAA	1260		
QY	1261	GAGAAAGCCTGGAGCCCTGGGATCGAGGTGAATCTACAGCCAAATGAGGAACTTGAGGCT	1320		
Db	1261	GAGAAAGCCTGGAGCCCTGGGATCGAGGTGAATCTACAGCCAAATGAGGAACTTGAGGCT	1320		
QY	1321	TTGGAAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATATAATGAAGAA	1380		
Db	1321	TTGGAAATGACGTATCTAATGGATGGGATCCCAATGATATGTTTCGATATAATGAAGAA	1380		
QY	1381	AAATTATGTTAGTGTCTACGTATCATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA	1440		
Db	1381	AAATTATGTTAGTGTCTACGTATCATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA	1440		
QY	1441	AGAGATAA	CTCAGAGAAATTTTTAAAA	CGGGAAGCAAGGGCAACACCGTTAGCAGAGAA	1500
Db	1441	AGAGATAA	CTCAGAGAAATTTTTAAAA	CGGGAAGCAAGGGCAACACCGTTAGCAGAGAA	1500
QY	1501	ATTGAGTCAAGTGCCTCAGTACAAAGCTCGAGTGGCCCTGGAAATGATAGAGTACAG	1560		
Db	1501	ATTGAGTCAAGTGCCTCAGTACAAAGCTCGAGTGGCCCTGGAAATGATAGAGTACAG	1560		
QY	1561	GAAAGAAAATACACAGCAGTTTCAAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAAC	1620		
Db	1561	GAAAGAAAATACACAGCAGTTTCAAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAAC	1620		
QY	1621	ACTAGGGAAAATAAATATATATTCCTCTGGACAAAGAAATAGAGAGTCAATTCCTGGGGA	1680		
Db	1621	ACTAGGGAAAATAAATATATATTCCTCTGGACAAAGAAATAGAGAGTCAATTCCTGGGGA	1680		
QY	1681	AGTGGGACACAGAAATTCACCGGTATGGCCAGCTCGATCGGGCTCCATGCCATCAAGA	1740		
Db	1681	AGTGGGACACAGAAATTCACCGGTATGGCCAGCTCGATCGGGCTCCATGCCATCAAGA	1740		
QY	1741	TCCACTTCTCACATTCAGATTTCAACCCGAAATTCGGTTCAGACCAAAGAGTAGTTAAT	1800		
Db	1741	TCCACTTCTCACATTCAGATTTCAACCCGAAATTCGGTTCAGACCAAAGAGTAGTTAAT	1800		
QY	1801	GGAGTGTTCCTGGCCATTCGCTTGGCCATCTCCTCTCGGCCACCTTCTCGCTAC	1860		
Db	1801	GGAGTGTTCCTGGCCATTCGCTTGGCCATCTCCTCTCGGCCACCTTCTCGCTAC	1860		
QY	1861	CAGTCAAGTCCCACTCTTCCACCTCGGCGAGCCACCTTACACGGCGCCCTCCAGG	1920		
Db	1861	CAGTCAAGTCCCACTCTTCCACCTCGGCGAGCCACCTTACACGGCGCCCTCCAGG	1920		
QY	1921	CCCCCTCGCGCCATCAGACCCCGCTCTACCCCTCTGCTCAATGGTTCTCCAGTCTCT	1980		
Db	1921	CCCCCTCGCGCCATCAGACCCCGCTCTACCCCTCTGCTCAATGGTTCTCCAGTCTCT	1980		

QY	1981	GTCTCTACTATGCTTAAACGCAATGTCTTTCAGAAAGGGCTTCCAAAGATGTCTCCCAAGGCC	2040			
Db	1981	GTCTCTACTATGCTTAAACGCAATGTCTTTCAGAAAGGGCTTCCAAAGATGTCTCCCAAGGCC	2040			
QY	2041	CAGCGACATCTCTCGAAATTCACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCTTA	2100			
Db	2041	CAGCGACATCTCTCGAAATTCACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCTTA	2100			
QY	2101	GAAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGACCACT	2160			
Db	2101	GAAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGACCACT	2160			
QY	2161	CCCTCGGGGGGAA	CGTGGTCAATCAGTGGGTTCAGAGTTATCCCTTAAAACT	2220		
Db	2161	CCCTCGGGGGGAA	CGTGGTCAATCAGTGGGTTCAGAGTTATCCCTTAAAACT	2220		
QY	2221	CATAGACCCAGGTCTCCAGACAGAA	CAGTATTCGAAATACCCCACTGGGCGAGTTCTT	2280		
Db	2221	CATAGACCCAGGTCTCCAGACAGAA	CAGTATTCGAAATACCCCACTGGGCGAGTTCTT	2280		
QY	2281	GCTTCTCCCAAGCTGTATTTCCAACTGAAGCTGTTCATGCTCTATTTCCAGCTGCA	2340			
Db	2281	GCTTCTCCCAAGCTGTATTTCCAACTGAAGCTGTTCATGCTCTATTTCCAGCTGCA	2340			
QY	2341	TCTCCTACGCTGTAGTCTCGCATTCGAA	CAGAGCTGTACCCCTTCTAGTGAGCTAAA	2400		
Db	2341	TCTCCTACGCTGTAGTCTCGCATTCGAA	CAGAGCTGTGTACCCCTTCTAGTGAGCTAAA	2400		
QY	2401	GATTTCCAGGCTTCAAGATCAGAGCGAC	CACTCTCCTCGAGGAAATTAAGAAATATTTAAA	2460		
Db	2401	GATTTCCAGGCTTCAAGATCAGAGCGAC	CACTCTCCTCGAGGAAATTAAGAAATATTTAAA	2460		
QY	2461	CCCAATGAAA	CATCACCTAGCTTCTCAAAAGCTGAAAACAAAGGTATATACACAGTTGTT	2520		
Db	2461	CCCAATGAAA	CATCACCTAGCTTCTCAAAAGCTGAAAACAAAGGTATATACACAGTTGTT	2520		
QY	2521	TCTGAA	CATAGAAAA	CAGATTTGATGATTTTAAAGAAA	TTTAAAGAAATTTTAAAGTATGATTTAGTTACAG	2580
Db	2521	TCTGAA	CATAGAAAA	CAGATTTGATGATTTTAAAGAAA	TTTAAAGAAATTTTAAAGTATGATTTAGTTACAG	2580
QY	2581	CCAAGTCTACTTCTGATCAAGACAA	AAATTTGAACCAAGTCTTCAAGGATTTCTTTCAATGAAAAT	2640		
Db	2581	CCAAGTCTACTTCTGATCAAGACAA	AAATTTGAACCAAGTCTTCAAGGATTTCTTTCAATGAAAAT	2640		
QY	2641	TCAAGAGATTTGATCAAGACAA	AAATTTGAACCAAGTCTTCAAGGATTTCTTTCAATGAAAAT	2700		
Db	2641	TCAAGAGATTTGATCAAGACAA	AAATTTGAACCAAGTCTTCAAGGATTTCTTTCAATGAAAAT	2700		
QY	2701	AGCAGCAGCAACTGTACAGTGGCAGCAG	CAGCAGCAAGTCTTCCCTTCA	2760		
Db	2701	AGCAGCAGCAACTGTACAGTGGCAGCAG	CAGCAGCAAGTCTTCCCTTCA	2760		
QY	2761	ATACCTTAGTAA	CAGGAGCAAGGAGGACCTCAGGTCTCTCCAGGGGTTTCAGACT	2820		
Db	2761	ATACCTTAGTAA	CAGGAGCAAGGAGGACCTCAGGTCTCTCCAGGGGTTTCAGACT	2820		
QY	2821	TCCAGCCAGCATGTAAACAAAGAA	AGACGATTAAGGAAGAAAGACGAGCTGAG	2880		
Db	2821	TCCAGCCAGCATGTAAACAAAGAA	AGACGATTAAGGAAGAAAGACGAGCTGAG	2880		
QY	2881	CAAGTTAGGAAATCAACATTTGAA	TCCCAATGCAAGGAGTTCAACCCACGTTCTCTCT	2940		
Db	2881	CAAGTTAGGAAATCAACATTTGAA	TCCCAATGCAAGGAGTTCAACCCACGTTCTCTCT	2940		
QY	2941	CAGCCAAAGCCTTCTACTACCCCA	AACTTCACTCGGCTCAGCACAACCTTAGCCCATCT	3000		
Db	2941	CAGCCAAAGCCTTCTACTACCCCA	AACTTCACTCGGCTCAGCACAACCTTAGCCCATCT	3000		
QY	3001	ATGGTGGGTCA	TCAACAGCCAACTCCAGTTTATATCTCAGCCTGTTTGTGTTTGCACCAAT	3060		
Db	3001	ATGGTGGGTCA	TCAACAGCCAACTCCAGTTTATATCTCAGCCTGTTTGTGTTTGCACCAAT	3060		
QY	3061	ATGATGATTC	AGTCCAGTGAGCCAGCGGTGCAACCTTTATACCCAATACCTATGACG	3120		

|||||
3061 ATGATGATTCAGTCCAGTCCAGTGAGCCAGCGGTGCAACCTTTATACCCAAATACCTATGAGC 3120
|||||
3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCCAAACAGCGG 3180
|||||
3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCCAAACAGCGG 3180
|||||
3181 CAAGACCAGCATCATCAGAGTGCCTATGATGACCCAGCGTTCAGCAGCGCGGCCACCGATT 3240
|||||
3181 CAAGACCAGCATCATCAGAGTGCCTATGATGACCCAGCGTTCAGCAGCGCGGCCACCGATT 3240
|||||
3241 GCAGCCACCCACCCAGCTTACTCCAGCAATATGTTGCCCTACAGTCCTCAGCAGTTCCTCA 3300
|||||
3241 GCAGCCACCCACCCAGCTTACTCCAGCAATATGTTGCCCTACAGTCCTCAGCAGTTCCTCA 3300
|||||
3301 AATCAGCCCTTTGTTCCAGCATGTGCAATATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
|||||
3301 AATCAGCCCTTTGTTCCAGCATGTGCAATATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
|||||
3361 CTTGTAATACAGGGTAACTGTAAGATGATGGCAACCAACACAGCCCGAGCCTGGTTTA 3420
|||||
3361 CTTGTAATACAGGGTAACTGTAAGATGATGGCAACCAACACAGCCCGAGCCTGGTTTA 3420
|||||
3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCATGCGATGTATGATGT 3480
|||||
3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCATGCGATGTATGATGT 3480
|||||
3481 CCCAAATACCATACAAAGGAGCAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3540
|||||
3481 CCCAAATACCATACAAAGGAGCAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3540
|||||
3541 TCCCTTGCTCAGCAGTATGGCAACCTTAACGCTACCCCTGCAACCAATACCTCAGACCCCT 3600
|||||
3541 TCCCTTGCTCAGCAGTATGGCAACCTTAACGCTACCCCTGCAACCAATACCTCAGACCCCT 3600
|||||
3601 CAGCTTCAGTACCCCTGAGCAGCAGCAAGCAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3660
|||||
3601 CAGCTTCAGTACCCCTGAGCAGCAGCAAGCAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3660
|||||
3661 CCAGTCTCTGTTCCAGCACCATCAGCAGCAGCCCGCCAGGCTCTCCATCTGGCCAGTCCA 3720
|||||
3661 CCAGTCTCTGTTCCAGCACCATCAGCAGCAGCCCGCCAGGCTCTCCATCTGGCCAGTCCA 3720
|||||
3721 CAGCAGAGTCAGCCATTTACACGGGGGCTTGGCCCAACTCCACCTTCCATGACACCT 3780
|||||
3721 CAGCAGAGTCAGCCATTTACACGGGGGCTTGGCCCAACTCCACCTTCCATGACACCT 3780
|||||
3781 GCCTCCAAACAGCAGTCGCCACAGATAGTTTCCAGCAGCAACACAGACTGCTTTTACG 3840
|||||
3781 GCCTCCAAACAGCAGTCGCCACAGATAGTTTCCAGCAGCAACACAGACTGCTTTTACG 3840
|||||
3841 ATCCATCTCTTCCAGCTTCCAGCCGGGTATACCAACCCACCCACATGGCCCACTACCT 3900
|||||
3841 ATCCATCTCTTCCAGCTTCCAGCCGGGTATACCAACCCACCCACATGGCCCACTACCT 3900
|||||
3901 CAGGCTCATGTA CAGTCAGGAATGGTTCTTCTTCTATCCAACTGCCCATGGCCCAATGATG 3960
|||||
3901 CAGGCTCATGTA CAGTCAGGAATGGTTCTTCTTCTATCCAACTGCCCATGGCCCAATGATG 3960
|||||
3961 CTATGACGACACAGCACCAGCGGGTCCCGAGCGCCCTCCCTCAAGTGCATACAG 4020
|||||
3961 CTATGACGACACAGCACCAGCGGGTCCCGAGCGCCCTCCCTCAAGTGCATACAG 4020
|||||
4021 CCCATTTCCAGTCTCGACAAACAGCGCATTTTCCCTTATATGACGACCTTTCAGTACAGCC 4080
|||||
4021 CCCATTTCCAGTCTCGACAAACAGCGCATTTTCCCTTATATGACGACCTTTCAGTACAGCC 4080
|||||
4081 CACCAACAAACAGCAGTTGTAAGGCTGCCTGGAGGAACCGAAAGGCCAAATTCCTTCCTC 4140
|||||
4081 CACCAACAAACAGCAGTTGTAAGGCTGCCTGGAGGAACCGAAAGGCCAAATTCCTTCCTC 4140
|||||
4141 CCTTCTACTGCTTCTACCAACTGGAGCAGAGAACTAGAAATTCATTTATTTTGT 4200

Db 4141 CCTTCTACTGCTTCTACCAACTGGAAGCACAGAAAACTAGAATTTCAATTTTGT 4200
Qy 4201 TAAATATATATGTTGATTTTCTGTAAACATCCATAGGAATGCTAACAGTTCACTTGCAG 4260
Db 4201 TAAATATATATGTTGATTTTCTGTAAACATCCATAGGAATGCTAACAGTTCACTTGCAG 4260
Qy 4261 TGGAAAGATACTTGGACCCAGTAGAGGCATTTAGGAACCTTGGGGGCTATTTCCTATAATTCCA 4320
Db 4261 TGGAAAGATACTTGGACCCAGTAGAGGCATTTAGGAACCTTGGGGGCTATTTCCTATAATTCCA 4320
Qy 4321 TATGCTGTTT CAGAGTCCCGCAGGTACCCCGAGCTCTGTTGCCGAACTGGGAAGTTATTT 4380
Db 4321 TATGCTGTTT CAGAGTCCCGCAGGTACCCCGAGCTCTGTTGCCGAACTGGGAAGTTATTT 4380
Qy 4381 ATTTTAAATACCCCTTGAAGTCTATGAACATCAGCTAGCAAAAGAGTAACAAGACT 4440
Db 4381 ATTTTAAATACCCCTTGAAGTCTATGAACATCAGCTAGCAAAAGAGTAACAAGACT 4440
Qy 4441 GATTTCTGCTGCTATTACTGCTTAAAAAAA 4481
Db 4441 GATTTCTGCTGCTATTACTGCTTAAAAAAA 4481
RESULT 7
CQ727533 4479 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 13467 from Patent WO02068579.
DEFINITION CQ727533
ACCESSION CQ727533
VERSION CQ727533.1 GI:42293778
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 13467 06-SEP-2002;
PE Corporation (NV) (US)
FEATURES
source 1..4479
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.3%; Score 4450.6; DB 6; Length 4479;..
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4475; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
Qy 1 ACCCCGAGAAAGCAACCCAGCGCGCCGCTCTCTCACGTGTCCTCCCGCCCGGG 60
Db 1 ACCCCGAGAAAGCAACCCAGCGCGCCGCTCTCTCACGTGTCCTCCCGCCCGGG 60
Qy 61 GCCACCTCAGTTCTGCTTCCGTCTGACCCCTCCGACTTCCGCTAAAGAGTCCCTATCCG 120
Db 61 GCCACCTCAGTTCTGCTTCCGTCTGACCCCTCCGACTTCCGCTAAAGAGTCCCTATCCG 120
Qy 121 CACTTCGGTCCGACCCCGCGGCTTGGCGGCGCCCTCCGATGCGCTCAGCGGCGCA 180
Db 121 CACTTCGGTCCGACCCCGCGGCTTGGCGGCGCCCTCCGATGCGCTCAGCGGCGCA 180
Qy 181 GCTTCCTCGAGTCCCGCGGTGGCCACCGAGTCTTCGCGCTTCGCCCGCAGCAGTGGCCC 240
Db 181 GCTTCCTCGAGTCCCGCGGTGGCCACCGAGTCTTCGCGG-TTCGCCCGCAGCAGTGGCCC 239
Qy 241 GGGTGGCGCTCGTCCAGCGGCGCGCGCGAGCGGGCGGGCGCGCGTGGCGCGCC 300
Db 240 GGGTGGCGCTCGTCCAGCGGCGCGCGCGAGCGGGCGGGCGCGCGTGGCGCGCC 299

QY 301 CCGGACCGTATCCCTCCGCGCCCTCCCGCGCCGCGCCCGCCCTCCCTCCCGG 360
Db |||||
QY 300 CCGGACCGTATCCCTCCGCGCCCTCCCGCGCCGCGCCCGCCCTCCCTCCCGG 359
Db |||||
QY 361 CAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGCGCGCG 420
Db |||||
QY 360 CAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGCGCGCG 419
QY 421 TTTCCGCGCGGCTCCGCGGCGCTCTTGGTTCGCGGCGCTCCCGCGCCCTTCGTCGTC 480
Db |||||
QY 420 TTTCCGCGCGGCTCCGCGGCGCTCTTGGTTCGCGGCGCTCCCGCGCCCTTCGTCGTC 479
QY 481 GTCTTCTCCCTCCGCGCGCGCGCTCCGCGGCTCCGCGGCGCGCGCGCGCGCGCTCCCGG 540
Db |||||
QY 480 CTCCTTCTCCCTCCGCGCGCGCGCTCCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCTCCCGG 539
QY 541 CTCGCGCGCGGCTCCGCGCGCGCTCCGCGGCTCCGCGGCTCCGCGGCGCGCGCGCGCGCG 600
Db |||||
QY 540 CTCGCGCGCGGCTCCGCGCGCGCTCCGCGGCTCCGCGGCTCCGCGGCGCGCGCGCGCGCG 599
QY 601 TGTCCCGCGCGCTCCGCGGCGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGG 660
Db |||||
QY 600 TGTCCCGCGCGCTCCGCGGCGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGG 659
QY 661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Db |||||
QY 660 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 719
QY 721 CAGCCCGCGCGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCT 780
Db |||||
QY 720 CAGCCCGCGCGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCT 779
QY 781 CCGCGCGCGCGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCT 840
Db |||||
QY 780 CCGCGCGCGCGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCT 838
QY 841 TCCTCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGAG 900
Db |||||
QY 839 TCCTCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGCGGCTCCGAG 898
QY 901 AGTAAACAAAGGACTCGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAAAATATGAGG 960
Db |||||
QY 899 AGTAAACAAAGGACTCGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAAAATATGAGG 958
QY 961 ATGGTTCATATCTACTACATCAGTCTTGGCTCCAAATGTGAAGTACAAAGTGAATAATGGA 1020
Db |||||
QY 959 ATGGTTCATATCTACTACATCAGTCTTGGCTCCAAATGTGAAGTACAAAGTGAATAATGGA 1018
QY 1021 GGTATATATGAGGAGTTTAAACTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1080
Db |||||
QY 1019 GGTATATATGAGGAGTTTAAACTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1078
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTTCGGGCGGAAACGTGAAGAAATATGAGAGGT 1140
Db |||||
QY 1079 GCACATGAGAAAGTACAGAAATCCAGTTTCGGGCGGAAACGTGAAGAAATATGAGAGGT 1138
QY 1141 ATTTTGTTCAAATGTTCAGACTTTGTTGGTACAGTTTAAAGATATGAGACTCCAGTTAT 1200
Db |||||
QY 1139 ATTTTGTTCAAATGTTCAGACTTTGTTGGTACAGTTTAAAGATATGAGACTCCAGTTAT 1198
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGAAGTGAATGAGGCGAACACAAA 1260
Db |||||
QY 1199 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGAAGTGAATGAGGCGAACACAAA 1258
QY 1261 GAGAAAGACCTTGGAGCCTGGGATCAGGTGAATCAGAGCAATGAGCACTTGAGGCT 1320
Db |||||
QY 1259 GAGAAAGACCTTGGAGCCTGGGATCAGGTGAATCAGAGCAATGAGCACTTGAGGCT 1318
QY 1321 TTGAAAAATGAGTATCTAATGAGTGGGATCCAAATGATATGTTTCGATATAATGAAGAA 1380
Db |||||
QY 1319 TTGAAAAATGAGTATCTAATGAGTGGGATCCAAATGATATGTTTCGATATAATGAAGAA 1378
QY 1381 AATTATGTTAGTGTCTACGTAATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1440

Db 1379 AATTATGTTAGTGTCTACGTAATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1438
QY 1441 AGAGATAACTCAGAAGAAATTTTAAACCGGAAGCAAGGGCAAAACCACTTAGCAAGAA 1500
Db 1439 AGAGATAACTCAGAAGAAATTTTAAACCGGAAGCAAGGGCAAAACCACTTAGCAAGAA 1498
QY 1501 ATTGAGTCAAGTGCCTCAGTACAAAGCTCAGTGGCCCTGGAAATATGATAGGAGTGAG 1560
Db 1499 ATTGAGTCAAGTGCCTCAGTACAAAGCTCAGTGGCCCTGGAAATATGATAGGAGTGAG 1558
QY 1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGSCACAGATAAAC 1620
Db 1559 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGSCACAGATAAAC 1618
QY 1621 ACTAGGGAAAAATAATATATTCCTCTCGGACAAAGAAATAGAGAGTCATATCTCTGGGA 1680
Db 1619 ACTAGGGAAAAATAATATATTCCTCTCGGACAAAGAAATAGAGAGTCATATCTCTGGGA 1678
QY 1681 AGTGGGAGACAGAAATTCACCGCTATGGGCGAGCTTGGATCGGCTCCATGCCATCAAGA 1740
Db 1679 AGTGGGAGACAGAAATTCACCGCTATGGGCGAGCTTGGATCGGCTCCATGCCATCAAGA 1738
QY 1741 TCCACTTCTCACACTTCAGATTTCAAACCGAAATTCAGTGAACGTGAGGGAGTATTAAT 1800
Db 1739 TCCACTTCTCACACTTCAGATTTCAAACCGAAATTCAGTGAACGTGAGGGAGTATTAAT 1798
QY 1801 GGAGTGTTCCTCGCCATCGCTTGGCCATCTCTCTCTCGCCCACTTCTCGCTAC 1860
Db 1799 GGAGTGTTCCTCGCCATCGCTTGGCCATCTCTCTCTCGCCCACTTCTCGCTAC 1858
QY 1861 CAGTCAGTTCCT 1920
Db 1859 CAGTCAGTTCCT 1918
QY 1921 CCCCCCTCGCGGCGCATCCAGACCCCGCTCTCAGCCCTCTCTCTCTCTCTCTCTCTCTCT 1980
Db 1919 CCCCCCTCGCGGCGCATCCAGACCCCGCTCTCAGCCCTCTCTCTCTCTCTCTCTCTCTCT 1978
QY 1981 GTCTCTACTATGCCCTAAACGATGTCTTTCAGAAAGGCGCTCCAAAGGATGTCTCCAAAGGCC 2040
Db 1979 GTCTCTACTATGCCCTAAACGATGTCTTTCAGAAAGGCGCTCCAAAGATGTCTCCAAAGGCC 2038
QY 2041 CAGCGACATCTCTCGAAATCAGAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Db 2039 CAGCGACATCTCTCGAAATCAGAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2098
QY 2101 GAATTTGTATCCCAACACCCACCGAGTGAAGCAGCTACTCTCTCTCTCTCTCTCTCTCTCT 2160
Db 2099 GAATTTGTATCCCAACACCCACCGAGTGAAGCAGCTACTCTCTCTCTCTCTCTCTCTCTCT 2158
QY 2161 CCCCCTCGGGGAAAGTGTCT 2220
Db 2159 CCCCCTCGGGGAAAGTGTCT 2218
QY 2221 CATAGACCCAGGTCTCCAGACAGAAAGATTTGGAATATCCCCAGTGGGCCAGTCTCT 2280
Db 2219 CATAGACCCAGGTCTCCAGACAGAAAGTATTTGGAATATCCCCAGTGGGCCAGTCTCT 2278
QY 2281 GCTTCTCCCAAGCTGGTATTTATTTCCAACTGAAGCTGTGTGCGCATGCTCTATTTCCAGCTGCA 2340
Db 2279 GCTTCTCCCAAGCTGGTATTTATTTCCAACTGAAGCTGTGTGCGCATGCTCTATTTCCAGCTGCA 2338
QY 2341 TCTCTCTACGCTCTAGTCTCTGCAACAGAGCTGTGTACCCCTCTTAGTGAGGCTAAA 2400
Db 2339 TCTCTCTACGCTCTAGTCTCTGCAACAGAGCTGTGTACCCCTCTTAGTGAGGCTAAA 2398
QY 2401 GATTCCAGGCTTCAAGATCAGAGCGAGAACTCTCTCTCGAGGGAATAAGAAAAATTTAAA 2460
Db 2399 GATTCCAGGCTTCAAGATCAGAGCGAGAACTCTCTCTCGAGGGAATAAGAAAAATTTAAA 2458
QY 2461 CCCAATGAACATCACCTAGCTCTCTCAAAAGCTGAAGCAAGGATATATCATCCAGTCTCT 2520

Db 2459 CCCAATGAAACATACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCGATTGTT 2518
Qy 2521 TCTGAACATAGAAAAACAGATTGATGATTTAAAGAAATTTAAGAAATGATTTTAGGTACAG 2580
Db 2519 TCTGAACATAGAAAAACAGATTGATGATTTAAAGAAATTTAAGAAATGATTTTAGGTACAG 2578
Qy 2581 CCAAGTTCTACTTCTGAAATCTATGATCTATGATCTAATACAAAATAGAGAGGAGAAAAA 2640
Db 2579 CCAAGTTCTACTTCTGAAATCTATGATCACTACTTAAACAAAATAGAGAGGAGAAAAA 2638
Qy 2641 TCAAGAGATTTTCATCAAAAGACAAAATGGAACCAAGTGTCTAAGGATTTCTTTTCATTCGAAAT 2700
Db 2639 TCAGAGATTTTCATCAAGACAAAATGGAACCAAGTGTCTAAGGATTTCTTTTCATTCGAAAT 2698
Qy 2701 AGCAGCAGCAACTGTATCCAGTGGCAGCAAGCCGAATAGCCCGACGATTTTCCCTTCA 2760
Db 2699 AGCAGCAGCAACTGTATCCAGTGGCAGCAAGCCGAATAGCCCGACGATTTTCCCTTCA 2758
Qy 2761 ATACTTAGTAAACAGGAGCAAGAGGGGACCTGAGGTCACTTCCAGGGGTTTCAAGCT 2820
Db 2759 ATACTTAGTAAACAGGAGCAAGAGGGGACCTGAGGTCACTTCCAGGGGTTTCAAGCT 2818
Qy 2821 TCCAGCCAGCATGTAAACAGAGAAAGACGATAGGAAGAGAGAAAGACGACGCTGAG 2880
Db 2819 TCCAGCCAGCATGTAAACAGAGAAAGACGATAGGAAGAGAGAAAGACGACGCTGAG 2878
Qy 2881 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAAAGGAGTTCAACCCAGCTTCTCT 2940
Db 2879 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAAAGGAGTTCAACCCAGCTTCTCT 2938
Qy 2941 CAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCCTCAAGCACAACCTAGCCCATCT 3000
Db 2939 CAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCCTCAAGCACAACCTAGCCCATCT 2998
Qy 3001 ATGGTGGGTATCAACAGCAACCTCAGTTTATCTCAGCCTGTTTGTGTTTCACCAAT 3060
Db 2999 ATGGTGGGTATCAACAGCAACCTCAGTTTATCTCAGCCTGTTTGTGTTTCACCAAT 3058
Qy 3061 ATGATGTATCCAGTCCAGTGAGCCGAGCGGTGCAACCTTTATACCCATACCTATGAG 3120
Db 3059 ATGATGTATCCAGTCCAGTGAGCCGAGCGGTGCAACCTTTATACCCATACCTATGAG 3118
Qy 3121 CCCATGCCAGTGAATCAAGCAAGACATATAGAGCAGTACCAAAATATGCCCCAACAGCGG 3180
Db 3119 CCCATGCCAGTGAATCAAGCAAGACATATAGAGCAGTACCAAAATATGCCCCAACAGCGG 3178
Qy 3181 CAAGACCAAGCATCATCAGAGTCCATGATGACCCAGCGCTCAGCAGCGGGCCCAACCGATT 3240
Db 3179 CAAGACCAAGCATCATCAGAGTCCATGATGACCCAGCGCTCAGCAGCGGGCCCAACCGATT 3238
Qy 3241 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCCCTACAGTCTCAGCAGTTCCTCA 3300
Db 3239 GCAGCCACCCACAGCTTACTCCAGCAATATGTTGCCCTACAGTCTCAGCAGTTCCTCA 3298
Qy 3301 AATCAGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCCTCATGCTATAGT 3360
Db 3299 AATCAGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCCTCATGCTATAGT 3358
Qy 3361 CTTGTTAATACAGGTTAATGCTAGAATGATGGACCAACCAACACAGCCGAGCTGTTTA 3420
Db 3359 CTTGTTAATACAGGTTAATGCTAGAATGATGGACCAACCAACACAGCCGAGCTGTTTA 3418
Qy 3421 GTATCTTCTTTCAGCAACTCAGTACGGGCTCATGAGCAGCATGCGATGATGATGT 3480
Db 3419 GTATCTTCTTTCAGCAACTCAGTACGGGCTCATGAGCAGCATGCGATGATGATGT 3478
Qy 3481 CCCAAATACCATACCAAGAGACAAGCCCTTCTTTCTACTTTGCCATTTCCACGGGC 3540
Db 3479 CCCAAATACCATACCAAGAGACAAGCCCTTCTTTCTACTTTGCCATTTCCACGGGC 3538
Qy 3541 TCCCTTGTCTCAGCAGTATGGGCAACCTTAAGCTTACCTTGACCCACATACTTCCACACCT 3600
Db 3539 TCCCTTGTCTCAGCAGTATGGGCAACCTTAAGCTTACCTTGACCCACATACTTCCACACCT 3598

Qy 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAAAGCCAAACATGGTGAAGTCATCTCTCA 3660
Db 3599 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAAAGCCAAACATGGTGAAGTCATCTCTCA 3658
Qy 3661 CCCAGTCTCTGTTTACAGCAACATCAGCAGCAGCGGCCAGGCTCTCCATCTGGCCAGTCCA 3720
Db 3659 CCCAGTCTCTGTTTACAGCAACATCAGCAGCAGCGGCCAGGCTCTCCATCTGGCCAGTCCA 3718
Qy 3721 CAGCAGCAGTCAAGCATTTTACCAGCGGGGCTTGGCCAACTTCCACCTTCCATGACACT 3780
Db 3719 CAGCAGCAGTCAAGCATTTTACCAGCGGGGCTTGGCCAACTTCCACCTTCCATGACACT 3778
Qy 3781 GCCTCCAAACAGCAGTTCGCAACAGATAGTTTCCAGCAGCACAACAGACTGTCTTTAGC 3840
Db 3779 GCCTCCAAACAGCAGTTCGCAACAGATAGTTTCCAGCAGCACAACAGACTGTCTTTAGC 3838
Qy 3841 ATCCATCTCTTCTCAGCTTTCAGCGCGGTATACCAACCCACCCACATGCGCCACCTACT 3900
Db 3839 ATCCATCTCTTCTCAGCTTTCAGCGCGGTATACCAACCCACCCACATGCGCCACCTACT 3898
Qy 3901 CAGGCTCATGTACAGTCAAGGAATGGTTCTTCTCATCCAACCTGCCCATGCGCCAATGATG 3960
Db 3899 CAGGCTCATGTACAGTCAAGGAATGGTTCTTCTCATCCAACCTGCCCATGCGCCAATGATG 3958
Qy 3961 CTAATGACGACACAGCCACCCGCGGTCCTCCAGCGCGGCTCGCTCAAAGTGCACTACAG 4020
Db 3959 CTAATGACGACACAGCCACCCGCGGTCCTCCAGCGCGGCTCGCTCAAAGTGCACTACAG 4018
Qy 4021 CCCATTCAGTCTGACAAACAGCGCATTTCCCTATATAGCAGCCCTTTCAGTACAAGCC 4080
Db 4019 CCCATTCAGTCTGACAAACAGCGCATTTCCCTATATAGCAGCCCTTTCAGTACAAGCC 4078
Qy 4081 CACCAACCAACAGCAGTTGTAAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTTCCCTCTC 4140
Db 4079 CACCAACCAACAGCAGTTGTAAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTTCCCTCTC 4138
Qy 4141 CTTCTACTGCTTCTACAACTGGAAGCAGAGAAACTAGAAATTTTCATTTATTTTGT 4200
Db 4139 CTTCTACTGCTTCTACAACTGGAAGCAGAGAAACTAGAAATTTTCATTTATTTTGT 4198
Qy 4201 TAAATATATATGTTGATTTCTTTAAACATCAATAGGAATGCTAACAGTTTCACTTGCAG 4260
Db 4199 TAAATATATATGTTGATTTCTTTAAACATCAATAGGAATGCTAACAGTTTCACTTGCAG 4258
Qy 4261 TGAAGAATATCTGGACCGAGTAGAGGCAATTTAGGAATCTTGGGGGCTATTTCATAATTCCA 4320
Db 4259 TGAAGAATATCTGGACCGAGTAGAGGCAATTTAGGAATCTTGGGGGCTATTTCATAATTCCA 4318
Qy 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGGCGAACTGGGAAGTTATTT 4380
Db 4319 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGGCGAACTGGGAAGTTATTT 4378
Qy 4381 ATTTTAAATACCTTGAAGTTCATGACATCAGCTAGCAAAAGAGTAACAGAGT 4440
Db 4379 ATTTTAAATACCTTGAAGTTCATGACATCAGCTAGCAAAAGAGTAACAGAGT 4438
Qy 4441 GATTTCTGCTGCTATTACTGCTTAAAAAATAAAAAAAAAAAAAA 4481
Db 4439 GATTTCTGCTGCTATTACTGCTTAAAAAATAAAAAAAAAAAAAA 4479

RESULT 8

HSNANSCA2
LOCUS HSDANSCA2 4163 bp mRNA linear PRI 18-APR-2005
DEFINITION H.sapiens mRNA for SCA2 protein.
ACCESSION Y08262
VERSION Y08262.1 GI:1770389
KEYWORDS SCA2 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

[illegible]

QY 1781 CAGACAAAGAGTAGTTAATGAGGTGTTCTCTGCGCATCGCTTGGCCATCTCTTCT 1840
DB CAGACAAAGAGTAGTTAATGAGGTGTTCTCTGCGCATCGCTTGGCCATCTCTTCT 1440
QY 1841 CTCGCCACCTTCTCGCTACCACTCAGGTCCAACTCTCTTCCACCTCGGCGAGCCACC 1900
DB CTCGCCACCTTCTCGCTACCACTCAGGTCCAACTCTCTTCCACCTCGGCGAGCCACC 1500
QY 1901 CTACACGGCGCCCTCCAGGCCCCCTCGCGCCCATCCAGACCCCCTCTCACCCCTCTG 1960
DB CTACACGGCGCCCTCCAGGCCCCCTCGCGCCCATCCAGACCCCCTCTCACCCCTCTG 1560
QY 1961 CTCATGGTCTCCAGTCTCTGTCTCTATATGCTTAAAGCATGTCTTCAGAAAGGCGCTC 2020
DB CTCATGGTCTCCAGTCTCTGTCTCTATATGCTTAAAGCATGTCTTCAGAAAGGCGCTC 1620
QY 2021 CAAGGATGTCCTCAAGGCCAGCGACATCTCGAAATCACAGAGTTCTGCTGGAGGG 2080
DB CAAGGATGTCCTCAAGGCCAGCGACATCTCGAAATCACAGAGTTCTGCTGGAGGG 1680
QY 2081 GTTTCATATCCAGTGGCTTAGAATTTGTATCCAAACCCACCCAGTGAAGCAGTACTC 2140
DB GTTTCATATCCAGTGGCTTAGAATTTGTATCCAAACCCACCCAGTGAAGCAGTACTC 1740
QY 2141 CTCAGTAGCAAGGACCACTCCCTCGGGGGGAACTGTGTCTATCAGTGGTCACTGGGGTTC 2200
DB CTCAGTAGCAAGGACCACTCCCTCGGGGGGAACTGTGTCTATCAGTGGTCACTGGGGTTC 1800
QY 2201 CAAGATTATCCCTTAAACTCATAGACCAGGTCTCCAGACAGAAACAGTATTGAAATA 2260
DB CAAGATTATCCCTTAAACTCATAGACCAGGTCTCCAGACAGAAACAGTATTGAAATA 1860
QY 2261 CCCCAGTGGGCGAGTCTTGTCTCTCCCAAGCTGTATTTCTCAACTGAAGCTGTG 2320
DB CCCCAGTGGGCGAGTCTTGTCTCTCCCAAGCTGTATTTCTCAACTGAAGCTGTG 1920
QY 2321 CCATGCTATTCAGCTGATCTCTACGCTCTAGTCTGATCGAAGCAGAGCTGTTA 2380
DB CCATGCTATTCAGCTGATCTCTACGCTCTAGTCTGATCGAAGCAGAGCTGTTA 1980
QY 2381 CCCCCTTAGTGAGGTAAAGATTCAAGGTTCAGGTTCAGATCAGAGCGAAGTCTCTGAG 2440
DB CCCCCTTAGTGAGGTAAAGATTCAAGGTTCAGGTTCAGATCAGAGCGAAGTCTCTGAG 2040
QY 2441 GGAATAAGAAATATTAACCAATGAACATCACTAGCTTCTCAAAAGCTGAAACA 2500
DB GGAATAAGAAATATTAACCAATGAACATCACTAGCTTCTCAAAAGCTGAAACA 2100
QY 2501 AAGGTATATCACCAGTGTGTTCTGAAATAGAAACAGATTGATGATTTAAAGAAATTTA 2560
DB AAGGTATATCACCAGTGTGTTCTGAAATAGAAACAGATTGATGATTTAAAGAAATTTA 2160
QY 2561 AGAATGATTTAGTTTACGCCAAGTTCTACTTCTGAATCTATGGATCACTAATAA 2620
DB AGAATGATTTAGTTTACGCCAAGTTCTACTTCTGAATCTATGGATCACTAATAA 2220
QY 2621 AATAAGAGGGAGAAAATCAAGAGATTGATCAAGACAAAATGAACCAAGTGCTA 2680
DB AATAAGAGGGAGAAAATCAAGAGATTGATCAAGACAAAATGAACCAAGTGCTA 2280
QY 2681 AGGATTTCTTTCATTTAGAAATAGCAGCAACTGTACCACTGGCAGCAGCAAGCCGAA 2740
DB AGGATTTCTTTCATTTAGAAATAGCAGCAACTGTACCACTGGCAGCAGCAAGCCGAA 2340
QY 2741 GCCCAGCAATTTCCCTTCAATCTTAGTAAACCGGAGCAAGAGGGGACCTGAGGTCA 2800
DB GCCCAGCAATTTCCCTTCAATCTTAGTAAACCGGAGCAAGAGGGGACCTGAGGTCA 2400
QY 2801 CTTCCCAAGGGGTTCACTTCCAGCCAGCATGTAAACAGAGAAAGACGATTAAGGAAG 2860
DB CTTCCCAAGGGGTTCACTTCCAGCCAGCATGTAAACAGAGAAAGACGATTAAGGAAG 2460

QY 2861 AGAAGAGAGCGCAGTGGCAAGTTAGGAATCAACATTCGAATCCCAATCGAAAGGAGT 2920
DB AGAAGAGAGCGCAGTGGCAAGTTAGGAATCAACATTCGAATCCCAATCGAAAGGAGT 2520
QY 2921 TCAACCCACGTTCTCTCTCAGCCAAAGCCCTTCTACTACCCCAACTTTCACCTCGGCTC 2980
DB TCAACCCACGTTCTCTCTCAGCCAAAGCCCTTCTACTACCCCAACTTTCACCTCGGCTC 2580
QY 2981 AAGCAAACTAGCCATCTATGTTGGTCTATCAACAGCCAACTCCAGTTTATCTACAG 3040
DB AAGCAAACTAGCCATCTATGTTGGTCTATCAACAGCCAACTCCAGTTTATCTACAG 2640
QY 3041 CTGTTGTTTTCGACCAAAATATGATTCAGTCCCACTCAGTGGCCAGCGTGCACCTT 3100
DB CTGTTGTTTTCGACCAAAATATGATTCAGTCCCACTCAGTGGCCAGCGTGCACCTT 2696
QY 3101 TATACCCCAATCTATGAGCCCATGCGCAGTGAATCAAGCCCAAGACATATAGAGCAGTAC 3160
DB TATACCCCAATCTATGAGCCCATGCGCAGTGAATCAAGCCCAAGACATATAGAGCAGTAC 2699
QY 3161 CAAATATCCCAACAGCGGCAAGACAGCATCATCAGAGTGCATGATGCACCCAGCT 3220
DB CAAATATCCCAACAGCGGCAAGACAGCATCATCAGAGTGCATGATGCACCCAGCT 2759
QY 3221 CAGCAGCGGCGCCAGTTCGAGCCACCCACAGCTTACTCCACGCAATATGTTGCT 3280
DB CAGCAGCGGCGCCAGTTCGAGCCACCCACAGCTTACTCCACGCAATATGTTGCT 2819
QY 3281 ACAGTCTCAGCAGTTCCTCAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTC 3340
DB ACAGTCTCAGCAGTTCCTCAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTC 2879
QY 3341 AGCATCTCTCATGTCTATAGTCTCTGTAAATACAGGGTAACTGTAAGATGATGGCAACCAA 3400
DB AGCATCTCTCATGTCTATAGTCTCTGTAAATACAGGGTAACTGTAAGATGATGGCAACCAA 2939
QY 3401 CACAGCCGAGCTGGTTAGTATCTTTCAGCAACTCAGTACGGGCTCATGAGCAGA 3460
DB CACAGCCGAGCTGGTTAGTATCTTTCAGCAACTCAGTACGGGCTCATGAGCAGA 2999
QY 3461 CCGATGCGATGATCATGTCTGCAAAATACCAATCAAGAGAGCAAGCCCTTCTTTCT 3520
DB CCGATGCGATGATGATGATGCTCCAAATACCAATCAAGAGAGCAAGCCCTTCTTTCT 3059
QY 3521 ACTTTGCCATTTCCAGGGCTCTCTTGTCTCAGCAGTATGCGCACCTAAAGCTACCTGC 3580
DB ACTTTGCCATTTCCAGGGCTCTCTTGTCTCAGCAGTATGCGCACCTAAAGCTACCTGC 3119
QY 3581 ACCCACAATCTCCACACCTCTAGCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 3640
DB ACCCACAATCTCCACACCTCTAGCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 3179
QY 3641 ATGGTGGAGTATCTCTGCACTCAGTCTCTGTTTACAGCATCAGCAGGCGCCAGG 3700
DB ATGGTGGAGTATCTCTGCACTCAGTCTCTGTTTACAGCATCAGCAGGCGCCAGG 3239
QY 3701 CTCCTCATCTGCGCAGTCCACAGCAGCAGTCTAGCCATTTTACACGCGGCTTGGCCCAA 3760
DB CTCCTCATCTGCGCAGTCCACAGCAGCAGTCTAGCCATTTTACACGCGGCTTGGCCCAA 3299
QY 3761 CTCCACCTCTCATGACCTCTGCTCCACACGAGTCTGCGCACAGAAATAGTTTCCAGCAG 3820
DB CTCCACCTCTCATGACCTCTGCTCCACACGAGTCTGCGCACAGAAATAGTTTCCAGCAG 3359
QY 3821 CACACAGACTGTCTTTAGCATCTCTTCTCAGTTCAGGCTTACGCGGCTATACCAACCCAC 3880
DB CACACAGACTGTCTTTAGCATCTCTTCTCAGTTCAGGCTTACGCGGCTATACCAACCCAC 3419
QY 3881 CCCCATGCGCCAGTACCTCAGGCTCATGTACAGTCAAGAAATGTTCTTCTCATCCAA 3940
DB CCCCATGCGCCAGTACCTCAGGCTCATGTACAGTCAAGAAATGTTCTTCTCATCCAA 3479
QY 3941 CTGCCCATGCGCAATGATGCTAATGACGACACAGCCACCCGCGGCTTCCAGCGGCC 4000

QY 1541 AAAATGATGATAGGAGTGAAGGAGAAATAACAGCAGCTTCAGAGAAATTCAGTGAAC 1600
| | | | |
Db 1141 AAAATGATGATAGGAGTGAAGGAGAAATAACAGCAGCTTCAGAGAAATTCAGTGAAC 1200
| | | | |
QY 1601 GTGAGGGGCACAGCATAAACACTAGGGAATAATAATATATCTCTCGACAAAGAAATA 1660
| | | | |
Db 1201 GTGAGGGGCACAGCATAAACACTAGGGAATAATAATATATCTCTCGACAAAGAAATA 1260
| | | | |
QY 1661 GAGAGTGCATATCTCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGGCCAGCTGGAT 1720
| | | | |
Db 1261 GAGAGTGCATATCTCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGGCCAGCTGGAT 1320
| | | | |
QY 1721 CGGGCTCATGCAATCAAGATCCACTTCTCACATTCAGATTTCAACCCGAATTCCTGGTT 1780
| | | | |
Db 1321 CGGGCTCATGCAATCAAGATCCACTTCTCACATTCAGATTTCAACCCGAATTCCTGGTT 1380
| | | | |
QY 1781 CAGACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCTTGCCCATCTCTCTCT 1840
| | | | |
Db 1381 CAGACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCTTGCCCATCTCTCTCT 1440
| | | | |
QY 1841 CTGCCCCACCTTCTCGCTACCACTCAGGTCCCAACTCTCTTCCAGCTCGGGCAGCACCC 1900
| | | | |
Db 1441 CTGCCCCACCTTCTCGCTACCACTCAGGTCCCAACTCTCTTCCAGCTCGGGCAGCACCC 1500
| | | | |
QY 1901 CTACAGGGCGGCTTCAGGCCCTCGGGCCATCCAGACCCCGGTCTCACCCCTCTG 1960
| | | | |
Db 1501 CTAACAGGGCGGCTTCAGGCCCTCGGGCCATCCAGACCCCGGTCTCACCCCTCTG 1560
| | | | |
QY 1961 CTCATGGTCTCCAGCTCTCTCTCTACTATGCTAAACGCTGCTTCAGAGAGGCGCTC 2020
| | | | |
Db 1561 CTCATGGTCTCCAGCTCTCTCTCTACTATGCTAAACGCTGCTTCAGAGAGGCGCTC 1620
| | | | |
QY 2021 CAAGGATGTCCTCCAAAGGCGCAGCAGATCTCTCGAAATCACAGAGTTTCTGCTGGAGGG 2080
| | | | |
Db 1621 CAAGGATGTCCTCCAAAGGCGCAGCAGATCTCTCGAAATCACAGAGTTTCTGCTGGAGGG 1680
| | | | |
QY 2081 GTTCCATATCCAGTGGCTAGAAATTTGTATCCACAAACCCAGTGAAGCGAGTACTC 2140
| | | | |
Db 1681 GTTCCATATCCAGTGGCTAGAAATTTGTATCCACAAACCCAGTGAAGCGAGTACTC 1740
| | | | |
QY 2141 CTCACGTAGCAAGGACGCTCCCTCGGGGGACGTTGTCATCAGTGGTCACTGGGGTTC 2200
| | | | |
Db 1741 CTCACGTAGCAAGGACGCTCCCTCGGGGGACGTTGTCATCAGTGGTCACTGGGGTTC 1800
| | | | |
QY 2201 CAAGATTTATCCCTAAACCTCATAGACCCAGGTCTCCAGACAGAACAGTATTGGAATA 2260
| | | | |
Db 1801 CAAGATTTATCCCTAAACCTCATAGACCCAGGTCTCCAGACAGAACAGTATTGGAATA 1860
| | | | |
QY 2261 CCCCCAGTGGGCGAGTTCTTGCTTCTCCCAAGCTGGTATATTCCAACTGAAGCTGTTG 2320
| | | | |
Db 1861 CCCCCAGTGGGCGAGTTCTTGCTTCTCCCAAGCTGGTATATTCCAACTGAAGCTGTTG 1920
| | | | |
QY 2321 CCAATGCTATCCAGCTGCTATCTCTACGCTCTAGTCTGCTGATCGAACAGAGCTGTTA 2380
| | | | |
Db 1921 CCAATGCTATCCAGCTGCTATCTCTACGCTCTAGTCTGCTGATCGAACAGAGCTGTTA 1980
| | | | |
QY 2381 CCCCTTCTAGTGGGCTAAAGATTCAGGCTTCAAGATCAGAGGCGAGAACTCTCTGCGAG 2440
| | | | |
Db 1981 CCCCTTCTAGTGGGCTAAAGATTCAGGCTTCAAGATCAGAGGCGAGAACTCTCTGCGAG 2040
| | | | |
QY 2441 GGAATAAGAAATAATTAATAACCAATGAACATCATCACCTAGCTTCTCAAAAGCTGAAACA 2500
| | | | |
Db 2041 GGAATAAGAAATAATTAATAACCAATGAACATCATCACCTAGCTTCTCAAAAGCTGAAACA 2100
| | | | |
QY 2501 AAGGTATATCAAGTTGTTCTGACATAGAAACAGATTTGATTTAAAGAAATTTA 2560
| | | | |
Db 2101 AAGGTATATCAAGTTGTTCTGAAATAGAAACAGATTTGATTTAAAGAAATTTA 2160
| | | | |
QY 2561 AAGATGATTTTAGGTTACAGCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACA 2620
| | | | |
Db 2161 AAGATGATTTTAGGTTACAGCAAGTTCTACTTCTGNAATCTATGGATCAACTACTAAACA 2220
| | | | |
QY 2621 AAAATAGAGAGGAGAGAAATCAAGAGATTTGTATCAAGAGCAAAATTTGAACCAAGTGCTA 2680
| | | | |

Db 2221 AAAATAGAGAGGAGAGAAATCAAGAGATTTGTATCAAGAGCAAAATTTGAAACCAAGTGCTA 2280
| | | | |
QY 2681 AGGATTTCTTTCAATGAAATAGCAGCAGCAACTGTGTACAGTGGCAGCAGCAAGCCGAATA 2740
| | | | |
Db 2281 AGGATTTCTTTCAATGAAATAGCAGCAGCAACTGTGTACAGTGGCAGCAGCAAGCCGAATA 2340
| | | | |
QY 2741 GCCCCAGCATTTCCGCTTCAATACCTAGTAACAGGAGCAAGAGGGGACCTGAGGTCA 2800
| | | | |
Db 2341 GCCCCAGCATTTCCGCTTCAATACCTAGTAACAGGAGCAAGAGGGGACCTGAGGTCA 2400
| | | | |
QY 2801 CTTTCCAAAGGGTTCAGACTTCCAGCCAGCATGTATAACAGAGAAAGACGATAGAGAG 2860
| | | | |
Db 2401 CTTTCCAAAGGGTTCAGACTTCCAGCCAGCATGTATAACAGAGAAAGACGATAGAGAG 2460
| | | | |
QY 2861 AGAAGAAAGACGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAAAGGAT 2920
| | | | |
Db 2461 AGAAGAAAGACGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAAAGGAT 2520
| | | | |
QY 2921 TCACCCAGGTTCTCTCTCAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCCTC 2980
| | | | |
Db 2521 TCACCCAGGTTCTCTCTCAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCCTC 2580
| | | | |
QY 2981 AAGCAAACTAGCCCATCTATGTGGTGCATCAACAGCAAACTCCAGTTTATATCTCAGC 3040
| | | | |
Db 2581 AAGCAAACTAGCCCATCTATGTGGTGCATCAACAGCAAACTCCAGTTTATATCTCAGC 2640
| | | | |
QY 3041 CTGTTTGTGTTGCAACCAATATGATGTATCCAGTCCCAGTGAGCCGCGTGCAACCTT 3100
| | | | |
Db 2641 CTGTTTGTGTTGCAACCAATATGATGTATCCAGTCCCAGTGAGCCGCGTGCAA---- 2696
| | | | |
QY 3101 TATACCCAAATACCTATAGCCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTAC 3160
| | | | |
Db 2697 -----TAC 2699
| | | | |
QY 3161 CAAATATGCCCAACAGCGGCAAGACCAGCATCATCAGAGTCCCATGTATGCCACCCAGCT 3220
| | | | |
Db 2700 CAAATATGCCCAACAGCGGCAAGACCAGCATCATCAGAGTCCCATGTATGCCACCCAGCT 2759
| | | | |
QY 3221 CAGCAGCGGGCCACCGATTCGAGCCACCCCAACAGCTTACTCCAGCAATATGTTGCTC 3280
| | | | |
Db 2760 CAGCAGCGGGCCACCGATTCGAGCCACCCCAACAGCTTACTCCAGCAATATGTTGCTC 2819
| | | | |
QY 3281 AAGTCTCAGCAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCAATATCAGTCTC 3340
| | | | |
Db 2820 AAGTCTCAGCAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCAATATCAGTCTC 2879
| | | | |
QY 3341 AGCATCTCATCTATAGTCTGTAATACAGGGTAATGCTAGAAATGATGGCAACCAAC 3400
| | | | |
Db 2880 AGCATCTCATCTATAGTCTGTAATACAGGGTAATGCTAGAAATGATGGCAACCAAC 2939
| | | | |
QY 3401 CACAGCCCGAGCTGGTTTGTATCTTCTCAGCAACTCAGTACGGGGCTCATAGCAGA 3460
| | | | |
Db 2940 CACAGCCCGAGCTGGTTTGTATCTTCTCAGCAACTCAGTACGGGGCTCATAGCAGA 2999
| | | | |
QY 3461 CGCATGCGATGATGCAATGCCAAATTAACCATACAAAGAGAGCAAGCCCTTCTTTCT 3520
| | | | |
Db 3000 CGCATGCGATGATGCAATGCCAAATTAACCATACAAAGAGAGCAAGCCCTTCTTTCT 3059
| | | | |
QY 3521 ACTTTGCCATTTCCAGGGCTCCCTTGTCTCAGCATGTGCCACCCCTAACGCTACCTGC 3580
| | | | |
Db 3060 ACTTTGCCATTTCCAGGGCTCCCTTGTCTCAGCATGTGCCACCCCTAACGCTACCTGC 3119
| | | | |
QY 3581 ACCCATATCTCCACACCCCTCAGCTTCACTAGTACCCCACTGGACAGCAGCAAGCAAC 3640
| | | | |
Db 3120 ACCCATATCTCCACACCCCTCAGCTTCACTAGTACCCCACTGGACAGCAGCAAGCAAC 3179
| | | | |
QY 3641 ATGGTGGAGTCACTCTGCAACCCAGTCTGTTTACGACCAATCAGACCCAGGCCCGCAGG 3700
| | | | |
Db 3180 ATGGTGGAGTCACTCTGCAACCCAGTCTGTTTACGACCAATCAGACCCAGGCCCGCAGG 3239
| | | | |
QY 3701 CTCTCCATCTGGCAGTCCACAGCAGCTCAGCCATTTTACCACGGCGCTTGGCCCAA 3760
| | | | |


```
|||||
2272 GTTAAACCTCGTTCTTTCTCAGCAAGCCTTCTACTACCCCAACGTCACCTCGGCC 2331
|||||
2979 TCAAGCACAACCTAGCCCATCTATGGTGGGTGTCATCAACAGAGCAACCTCCAGTTTATACCTCA 3038
|||||
2332 TCAAGCACAACCCAGCCCATCTATGGTGGGTGTCATCAGAGCCAGCTCCAGTGTCACCTCA 2391
|||||
3039 GCCTGTTGTTTGGACAACAATATGATGTATCCAGTCCAGTCCAGTGAGCCAGCGGTGCACACC 3098
|||||
2392 GCCTGTGTGCTTCGCAACCAATATGATGTATCCCGTCCAGTGAGCCCGGGCGGTACAACC 2451
|||||
3099 TTTATACCAATACCTATGAGGCCCATGCCAGTGAATCAAGCCAGAGACATATAGAC--- 3155
|||||
2452 TTTATACCAATACCTATGAGGCCCATGCCAGTGAATCAAGCCAGAGACATATAGAGCAGG 2511
|||||
3156 ---AGTACCAATATGCCCAACAGCGGCAAGACAGCATCATCAGAGTGCATGATGCA 3212
|||||
2512 TAAAGTACCAATATGCCCAACAGCGAGACAGACCAACATCATCAAGCACCATGATGCA 2571
|||||
3213 CCCAGCGTCAGAGCGGGGCCACCGATTGCAAGCCACCCACAGCTTACTCOAGCAATA 3272
|||||
2572 CCCAGCCTCGCGGAGGGCCACCCCATCGTAGCCACCCCGCGCTTACTCCACTCAGTA 2631
|||||
3273 TGTTGCTTACAGTCTCAGCAGTTCCTCAAAATCAGCCCTTGTTCAGCATGTGCCACATTA 3332
|||||
2632 CGTTGCTTACAGCCCTCAGCAGTTCCTCAAAATCAGCCCTTGTTCAGCATGTGCCACATTA 2691
|||||
3333 TCAGTCTCAGCATCCTCATGTCTATAGTCCGTGAATACAGGGTAAATGCTAGAAATGATGGC 3392
|||||
2692 TCAGTCTCAGCATCCTCATGTGTACAGTCCGTGTATACAGGTAATGCGAGATGATGGC 2751
|||||
3393 ACCACCAACAGAGCCCGAGCCGTGTTAGTATCTTTTCAGCAACTCAGTACGGGGCTCA 3452
|||||
2752 ACCACGACACATGCTCAGCCCTGGTTAGTGTCTTTCTTTTCAGTGTCTCAGTTCGGGGCTCA 2811
|||||
3453 TGAGCAGAGCGATGCGATGTATGATGTGCCATGTCCTCAAAATACATACAGAGGACAGAGCC 3512
|||||
2812 CGAGCAGAGCGAGCGCATGTATGATGTGCCATGTCCTCAAAATACATACAGAGGACAGAGCC 2871
|||||
3513 TTCTTTTCTACTTTGTCATTTTCCACGGGCTCCCTGTCTCAGAGATGTGGCCACCTTAAGCC 3572
|||||
2872 TTCCTTTCTACTTTGTCATTTTCCACGGGCTCCCTGTCTCAGAGATGTGGCCACCTTAAGCC 2931
|||||
3573 TACCTTGCAACCAATATCTCCACACCTCTCAGCCTTTCAGTACCTCCCACTGGAACAGCAGA 3632
|||||
2932 CGCCTGTGATCCACATACTCCCATCTCAGCCTTTCCGGCCACTCCACCGGACAGCAGCA 2991
|||||
3633 AAGCCAAACATGTTGGAAGTTCATCTGCAACCCAGTCTCTTTCAGCACCATCAGCACCGGCC 3692
|||||
2992 AAGCAGCATGTTGGAAGTTCACCTTGCAACCCAGTCTCTTTCAGCACCATCAGCACCGGCC 3051
|||||
3693 CGCCAGGCTCTTCCATCTGCGCAGTCCACAGCAGCAGTTCAGCCTTACCAACCGGGGCT 3752
|||||
3052 TGCCAGGCTCTTTCATCTGCGCAGTCCACAGCAGCAGTTCGGCCATTTATCATCGGGGCT 3111
|||||
3753 TGGCCAACTCCACCTCCATGACACCTGCTCCAAACAGCAGTCCGACAGCAAGATAGTTT 3812
|||||
3112 GGCACCAACACACCTTCATGACACCTTCCTCTAATACACAGTCTCCACAGAGCAGTTT 3171
|||||
3813 CCCAGCAGCAACAGACTGTCTTTACGATCCATCTCTTTCAGTTCAGCCGGGTATAC 3872
|||||
3172 CCCAGCAGCAACAGCAGTCTTCCACCATCCACCTTCTCATGTTTCAGCCGGCATACAC 3231
|||||
3873 CAACCCACCCACATGCGCCACGTACCTCAGGCTCATGTACAGTCAAGGAATGGTTCCTTC 3932
|||||
3232 CACCCCAACCCACATGCGCCACGTACCTCAGGCTCATGTACAGTCAAGGAATGGTTCCTTC 3291
|||||
3933 TCATCCAACTGCGCCATGCGCCAAATGATGTATGACGACAGCCACCCCGCGGTCCCA 3992
|||||
3292 TCATCCAACTGCGCCATGCGCCAAATGATGTATGACGACAGCCACCC---CGGTCCCAA 3348
|||||
3993 GCGCGCCTCGCTCAAGTGCACTACAGCCCATTCAGTCTCGACAAAGCGCATTTCCC 4052
|||||
```

```
Db 3349 GGGCGCCCTCGCTCAAAAGTGCACTACAGCCCATTTCCAGTTTCGACAAACAGGCGCATTTCCC 3408
QY 4053 CTATATGACGCAACCTCTTTCAGTACAAAGCCCAACCAACAGCAGTGTGAAGCTGCGCCCTGG 4112
|||||
Db 3409 TTATATGAGCGCACCTTTCAGTACAAAGCCCAACCAACAGCAGTGTGAAGCTGCGCCCTGG 3468
|||||
QY 4113 AGAAACGGAAGGCAAAATTCCTCTCCCTTCTTACTGCTCTTACCAACTCGAAGCAGACAG 4172
|||||
Db 3469 AGAAACGGAAGGCAAAATTCCTCTCCCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3527
|||||
QY 4173 AAAAAGTAAATTTCTATTTTCTTTTAAATATATATATATATATATATATATATATATATATATAT 4232
|||||
Db 3528 AAAAAGTAAATTTCTATTTTCTTTT-----TAAAGATACACTGATTTTAACTCT 3581
|||||
QY 4233 AATAGGAATGCTAAACAGTTCACCTTGCAGTGAAGAT-ACITTTGGACCCGAGTAGAGGCATTT 4291
|||||
Db 3582 GATAGGAATGCTAAACAGTTCACCTTGCAGTGAAGATGTTTGGACCCGAGTAGAGGCATGT 3641
|||||
QY 4292 AGAACTTTGGGGGCTATTCATTAATTCATATGCTGTTTCAGAGTCCCGCAGGTACCCCA 4351
|||||
Db 3642 AGGACCTT-GTGGCTGTTCATAATTCATGCTGTTTGCAGGGTCTCTGCAAGTA-CCCA 3699
|||||
QY 4352 GCTCTGCTTCGGAACCTGGAAGTTATTTTAAATACCCCTTGAAGCTCATGAACA 4411
|||||
Db 3700 GCTCTGCTTCGGAACCTGGAAGTTATTTTAAATGCGCCTTGAGAGTTCATGAACA 3759
|||||
QY 4412 CATCAGCTTAGCAAAAGAAAGTAAACAAAGAGTGAATCTTCTGCT 4450
|||||
Db 3760 CATCAGCTAGCAACAGAAAGTAAACAAAGAGTGAATCTTCTGCT 3798
|||||
```

```
RESULT 12
AK095017
LOCUS Homo sapiens cdna FLJ37698 fis, clone BRHIP2015679, highly similar
to Human ataxin-2 (SCA2) mRNA.
DEFINITION AK095017
ACCESSION AK095017.1 GI:21754198
VERSION oligo capping; fis (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM
```

```
REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
```

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirao, K., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yoshida, M., Hota, T., Kusano, J., Kanehori, K., Takahashi, F., Ueda, H., Tanase, T., Nomura, Y., Togawa, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagi, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,

DEFINITION Xenopus laevis cDNA clone MGC:115230 IMAGE:5079818, complete cds.
ACCESSION BC097692
VERSION BC097692.1 GI:66910767
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 2558)
AUTHORS Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
REFERENCE 2 (bases 1 to 2558)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schectz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
CONSRTM Mammalian Gene Collection Program Team
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 3 (bases 1 to 2558)
AUTHORS Klein, S. and Gerhard, D.S.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2005) National Institutes of Health, Xenopus Genome Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4801, Rockville, MD 20892-7510, USA
REMARK NIH-MGC Project
COMMENT Contact: XGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 229 Row: h Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES
 source
 Location/Qualifiers
 1..2558
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="MGC:115230 IMAGE:5079818"
 /tissue type="Oocytes"

Db 899 CAGAGAACTGCAGTCACCGGAGGGGCATGGCCCCAACACTAGGACACAATAATATATT 958
QY 1642 CCTCTTGGACAAAGAAATAGAGAAGTCATATCCTGGGGAAGTGGGAGACAGAATTTCACCG 1701
Db 959 CCTCTTGGACAAAGAAACAGAGAAGTCCTATCCTGGGGAAGTGGGAGACAGAGCTCACCA 1018
QY 1702 CGTATGGGCCAGCCTGGATCGGGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGAT 1761
Db 1019 CGGATGGGCCAGCCTGGGCCAGGCTCCATGCCGTCAAGAGCTGCTTCTCACACTTCAGAT 1078
QY 1762 TTCAACCCGAATTCTGGTTTCAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGGCATCG 1821
Db 1079 TTCAACCCGAACGCTGGCTCAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGGCATCG 1138
QY 1822 CCTTGGCCATCTCCTTCTCTCGGCCACCTTCTCGCTACAGTCAGGTCCCAACTCTCTT 1881
Db 1139 CCTTGGCCATCTCCTTCTCTCGGCCACCTTCTCGCTACAGTCAGGTCCCAACTCTCTT 1198
QY 1882 CCACCTCGGGCAGCCACCCCTACACGGCGCCCTCC 1917
Db 1199 CCACCTCGGGCAGCCACCCCTACACGGCCTCGTGCC 1234

Search completed: December 23, 2005, 00:44:05
Job time : 21615 secs

This Page Blank (uspto)

Result No.	Score	Query			DB	ID	Description
		Match	Length	*			
1	4481	100.0	4481	2	AAV06552	Human SCA	
2	4481	100.0	4481	2	AAV23428	Human SCA	
3	4481	100.0	4481	10	ADD18753	Human di	
4	4481	100.0	4481	10	ADH69466	Human sp	
5	4481	100.0	4481	10	ABV75291	Human SCA	
6	4481	100.0	4481	14	ADZ49164	Inseulin	
7	4347.8	97.0	4367	2	AAV30270	Gene caus	
8	3949	88.1	4200	2	AAT78912	Spinocere	
9	3138.2	70.0	4225	10	ABV75290	Mouse SCA	
10	2957	66.0	3798	2	AAV06553	Mouse SCA	
11	1730.6	38.6	2006	3	AAF21628	Human bre	
12	935.2	20.9	1257	10	ABV75289	Mouse SCA	
13	633	14.1	4369	13	ADRO8336	Full len	
14	571.2	12.7	623	2	AAV17229	SCA2 gene	
15	497.4	11.1	516	2	AAV06551	SCA2 gene	
16	455.8	10.2	459	10	ADH69479	Human SCA	
17	441.2	9.8	704	6	ABQ36595	Oligonuc	
18	441.2	9.8	704	6	ABQ36594	Oligonuc	
19	437.6	9.8	485	9	ACH41056	Human fo	

FT	repeat_unit	676. .678		PS
FT		/*tag= m		XX
FT	repeat_unit	/note= "CAG repeat"		CC
FT		679. .681		CC
FT		/*tag= n		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		685. .687		CC
FT		/*tag= o		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		688. .690		CC
FT		/*tag= p		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		691. .693		CC
FT		/*tag= q		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		694. .696		CC
FT		/*tag= r		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		700. .702		CC
FT		/*tag= s		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		703. .705		CC
FT		/*tag= t		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		706. .708		CC
FT		/*tag= u		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		709. .711		CC
FT		/*tag= v		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		712. .714		CC
FT		/*tag= w		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		715. .717		CC
FT		/*tag= x		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		718. .720		CC
FT		/*tag= y		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT		721. .723		CC
FT		/*tag= z		CC
FT	repeat_unit	/note= "CAG repeat"		CC
FT	primer_bind	740. .757		CC
FT		/*tag= c		CC
FT	exon	/note= "primer SCA2-B binding site"		CC
FT		899. .900		CC
FT		/*tag= e		CC
FT	primer_bind	/note= "predicted splice site"		CC
FT		1070. .1091		CC
FT		/*tag= d		CC
FT		/note= "primer SCA2-14B binding site"		CC
FT	WO9742314-A1.			CC
PN				XX
DD	13-NOV-1997.			XX
XX				XX
XX	08-MAY-1997;	97WO-US007725.		XX
XX				XX
PR	08-MAY-1996;	96US-0017388P.		XX
PR	19-JUL-1996;	96US-0022207P.		XX
PR	08-OCT-1996;	96US-00727084.		XX
XX				XX
PA	(CBDA-) CEDARS SINAI MEDICAL CENT.			XX
XX				XX
PI	Pulst S;			XX
XX				XX
DR	WPI; 1998-086523/08.			XX
DR	P-PSDB; AAW33807.			XX
XX				XX
PT	Nucleic acids encoding human and mouse ataxin 2 - a product of the			XX
PT	spinocerebellar ataxia 2 gene, SCA2; useful in the diagnosis of ataxia			XX
PT	type 2.			XX
XX				XX

Claim 6; Page 52-58; 98pp; English.

This cDNA sequence corresponds to a novel SCA2 gene encoding a human spinocerebellar ataxin-2 (SCA2) polypeptide, designated ataxin-2 (see AAW33807). A trisomy 21 foetal brain cDNA library and an adult human frontal cortex cDNA library in lambda ZapII were screened with probes obtained by PCR amplification of plasmid AAP65122B (see AAV06551). PCR products were used to screen the human adult frontal cortex library, and 5' clones were obtained by RT-PCR of placental mRNAs. Overlapping clones was used to generate the composite 4481 bp sequence. Ataxin type 2 can be diagnosed by detecting a genomic or transcribed mRNA sequence in an individual having an expanded CAG repeat at a location corresponding to the CAG repeat region of the SCA2 gene. The presence of at least 13 CAG repeats above the normal level (22, occasionally 23, repeats) is indicative of SCA2. Primers (see AAT99640-41) amplifying at least this region are used for diagnosis. Also claimed are kits for detecting mutations at the SCA2 locus, antisense oligonucleotides, and transgenic animals useful for studying the physiological roles of ataxin-2 and its effect upon behaviour

Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

Query Match	100.0%;	Score 4481;	DB 2;	Length 4481;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 4481;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCAGCTGTCCCTCCCGGCGCGCGG	60	
Db	1	ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCAGCTGTCCCTCCCGGCGCGCGG	60	
Qy	61	GCCACCTCACGTTCTGCTTCGGTCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG	120	
Db	61	GCCACCTCACGTTCTGCTTCGGTCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG	120	
Qy	121	CACCTCCGCTCCACCCCGGCGCTCGGGGCGCGCGCTCGATGCGCTCAGCGCGCGCA	180	
Db	121	CACCTCCGCTCCACCCCGGCGCTCGGGGCGCGCGCTCGATGCGCTCAGCGCGCGCA	180	
Qy	181	GCTCTCGGAGTCCCGGCGTGCCACCGAGTCTCGCGCTTCGCGCAGCAGCTGGCGCC	240	
Db	181	GCTCTCGGAGTCCCGGCGTGCCACCGAGTCTCGCGCTTCGCGCAGCAGCTGGCGCC	240	
Qy	241	GGGTGGCGCTCGTTCACGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG	300	
Db	241	GGGTGGCGCTCGTTCACGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG	300	
Qy	301	CCGGAGCGGTATCCCTCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCG	360	
Db	301	CCGGAGCGGTATCCCTCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCG	360	
Qy	361	CAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGCAACGGCGGCGCG	420	
Db	361	CAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGCAACGGCGGCGCG	420	
Qy	421	TTTCGGCGCGCTCCCGCGGCTCTTTGGTCTCGGCGGCTCTCCCGCGCGCTTCGTCTC	480	
Db	421	TTTCGGCGCGCTCCCGCGGCTCTTTGGTCTCGGCGGCTCTCCCGCGCGCTTCGTCTC	480	
Qy	481	GTCTTCTCTCCCTTCGCGCAGCCCGGCGCGCTTCGGCGCGCGCGCGCGCGCGCG	540	
Db	481	GTCTTCTCTCCCTTCGCGCAGCCCGGCGCGCTTCGGCGCGCGCGCGCGCGCGCG	540	
Qy	541	CTCGGCGCGCTCGCTCCCGCGCGCTTCGGCGGTCTCTTTGGCGCGCGCGCGCGCG	600	
Db	541	CTCGGCGCGCTCGCTCCCGCGCGCTTCGGCGGTCTCTTTGGCGCGCGCGCGCGCG	600	
Qy	601	TGTCTCCCGCGCGCTCGAGCCCGCTTATGGCGCTTCAACATGTCTGCTGAAGCCCG	660	
Db	601	TGTCTCCCGCGCGCTCGAGCCCGCTTATGGCGCTTCAACATGTCTGCTGAAGCCCG	660	
Qy	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720	
Db	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720	

QY 721 CAGCGCCGCCCGCGCTGCCAATGTCCGCAAGCCCGCGGCGAGCGGCTTCTAGCGTCG 780
DB 721 CAGCGCCGCCCGCGCTGCCAATGTCCGCAAGCCCGCGGCGAGCGGCTTCTAGCGTCG 780
QY 781 CCCGCCCGCGCCCTTCCGCGCTCCCTGCTCCGCTCTCCGCTCTCCGCTCCGCTCCGCTCC 840
DB 781 CCCGCCCGCGCCCTTCCGCGCTCCCTGCTCCGCTCTCCGCTCTCCGCTCCGCTCCGCTCC 840
QY 841 TCCTCGGTGTCGCGCGACCTTCCGCGCGCGGAGGCCCGCGCTCCGCGCGAGGTCGAAAC 900
DB 841 TCCTCGGTGTCGCGCGACCTTCCGCGCGCGGAGGCCCGCGCTCCGCGCGAGGTCGAAAC 900
QY 901 AGTAACAAAGGACCTGCTCAGCTACGATTTCTTTGATGGAATCTATGCAAAATATGAGG 960
DB 901 AGTAACAAAGGACCTGCTCAGCTACGATTTCTTTGATGGAATCTATGCAAAATATGAGG 960
QY 961 ATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAATGGA 1020
DB 961 ATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAATGGA 1020
QY 1021 GGTATATATGAAGGAGTTTAAAACTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1080
DB 1021 GGTATATATGAAGGAGTTTAAAACTTACAGTCCGAAGTGTGATTTGGTACTTGATGCC 1080
QY 1081 GCACATGAGAAAGTACAGAACTCCAGTTCGGGGCGGAAACGTGAAGAAATATGAGAGT 1140
DB 1081 GCACATGAGAAAGTACAGAACTCCAGTTCGGGGCGGAAACGTGAAGAAATATGAGAGT 1140
QY 1141 ATTTTGTTCAAATGTTTCAACCTTGTGTTGGTACAGTTTAAAGATATGACCTCCAGTTAT 1200
DB 1141 ATTTTGTTCAAATGTTTCAACCTTGTGTTGGTACAGTTTAAAGATATGACCTCCAGTTAT 1200
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTAAAGTGAATGGCGGAAACACAAA 1260
DB 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTAAAGTGAATGGCGGAAACACAAA 1260
QY 1261 GAGAGGACCTGGAGCCCTGGAGTACAGTGAACCTCAGACCAATAGAGAACTTGAGGCT 1320
DB 1261 GAGAGGACCTGGAGCCCTGGAGTACAGTGAACCTCAGACCAATAGAGAACTTGAGGCT 1320
QY 1321 TTGGAAAATGACGTATCTAATGGATGGATCCCAATGATGTTTCGATATATGAGAA 1380
DB 1321 TTGGAAAATGACGTATCTAATGGATGGATCCCAATGATGTTTCGATATATGAGAA 1380
QY 1381 AATTATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1440
DB 1381 AATTATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGCCTTAGAA 1440
QY 1441 AGAGATAACTCAGAGAAATTTTAAAAACGGGAAGCAAGGCGAAACAGTTAGCAGAA 1500
DB 1441 AGAGATAACTCAGAGAAATTTTAAAAACGGGAAGCAAGGCGAAACAGTTAGCAGAA 1500
QY 1501 ATTGAGTCAAGTCCAGTACAAAGCTCCAGTGGCCCTCGAATAATGATAGGAGTGA 1560
DB 1501 ATTGAGTCAAGTCCAGTACAAAGCTCCAGTGGCCCTCGAATAATGATAGGAGTGA 1560
QY 1561 GAAAGAAAATACACAGAGTTTCAAGAAATTTCCAGTGAACGTGAGGGGCGACAGCATAAAC 1620
DB 1561 GAAAGAAAATACACAGAGTTTCAAGAAATTTCCAGTGAACGTGAGGGGCGACAGCATAAAC 1620
QY 1621 ACTAGGAAAATAAATATATTCCTCTCGGACAAAGAAATAGAGAGTCTATCTCTGGGA 1680
DB 1621 ACTAGGAAAATAAATATATTCCTCTCGGACAAAGAAATAGAGAGTCTATCTCTGGGA 1680
QY 1681 AGTGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCGGGCTCCATGCCATCAAG 1740
DB 1681 AGTGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCGGGCTCCATGCCATCAAG 1740
QY 1741 TCCACTTCTCACAATTCAGATTTCAACCCGAAATTCCTGGTTTCAACCAAGAGTGTAAAT 1800
DB 1741 TCCACTTCTCACAATTCAGATTTCAACCCGAAATTCCTGGTTTCAACCAAGAGTGTAAAT 1800

QY 1801 GGAGTGTTCCTCGCCATCGCCTTGCCCATCTCTCTCTCTCGCCACCTTCTCGCTAC 1860
DB 1801 GGAGTGTTCCTCGCCATCGCCTTGCCCATCTCTCTCTCTCGCCACCTTCTCGCTAC 1860
QY 1861 CAGTCAGTTCGCAACTCTCTTCCAGCTCGGGCAGCACCCCTTACACGGCGCCCTCCAGG 1920
DB 1861 CAGTCAGTTCGCAACTCTCTTCCAGCTCGGGCAGCACCCCTTACACGGCGCCCTCCAGG 1920
QY 1921 CCCCTCGCGGCGCATCCAGACCCCCTCTACCCCTCTGCTCATGGTCTCCAGTCTCT 1980
DB 1921 CCCCTCGCGGCGCATCCAGACCCCCTCTACCCCTCTGCTCATGGTCTCCAGTCTCT 1980
QY 1981 GTCTCTATGTCCTTAAACGCACTGTTCAGAAAGGCGCTTCCAAAGGATGTCCCAAAGGCC 2040
DB 1981 GTCTCTATGTCCTTAAACGCACTGTTCAGAAAGGCGCTTCCAAAGGATGTCCCAAAGGCC 2040
QY 2041 CAGCGACATCTCTCGAAATCAGAGTTCCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100
DB 2041 CAGCGACATCTCTCGAAATCAGAGTTCCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100
QY 2101 GAATTTGTATCCCAACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGGACCACT 2160
DB 2101 GAATTTGTATCCCAACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGGACCACT 2160
QY 2161 CCCTCGGGGGAAACGTGTGTCATCAGTGTGTCAGTGGGTTTCCAAAGATTTATCCCTTAAACT 2220
DB 2161 CCCTCGGGGGAAACGTGTGTCATCAGTGTGTCAGTGGGTTTCCAAAGATTTATCCCTTAAACT 2220
QY 2221 CATAGACCCAGTCTCCAGACAGAACAGTATTTGAAATATACCCAGTGGGCGAGTTCTT 2280
DB 2221 CATAGACCCAGTCTCCAGACAGAACAGTATTTGAAATATACCCAGTGGGCGAGTTCTT 2280
QY 2281 GCTTCTCCCAAGCTGTTATTTCAACTGAAGCTGTTCAGTGGGTTTCCAAAGATTTATCCCTTAAACT 2340
DB 2281 GCTTCTCCCAAGCTGTTATTTCAACTGAAGCTGTTCAGTGGGTTTCCAAAGATTTATCCCTTAAACT 2340
QY 2341 TCTCTACGCTGTAGTCTCGATCGAACAGCTGTTCACCCCTTCTAGTCAGGCTAA 2400
DB 2341 TCTCTACGCTGTAGTCTCGATCGAACAGCTGTTCACCCCTTCTAGTCAGGCTAA 2400
QY 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAAGAAATATATAA 2460
DB 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAAGAAATATATAA 2460
QY 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTAATATCAACAGTTGTT 2520
DB 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTAATATCAACAGTTGTT 2520
QY 2521 TCTGAACATAGAAACAGATTCGATTTAAAGAAATTTAAGAAATGATTTTAGGTTACAG 2580
DB 2521 TCTGAACATAGAAACAGATTCGATTTAAAGAAATTTAAGAAATGATTTTAGGTTACAG 2580
QY 2581 CCAAGTTCTTCTTGAATCTATGATCAATCTAAACAAATAAGAGGGAGAGAAAA 2640
DB 2581 CCAAGTTCTTCTTGAATCTATGATCAATCTTAAACAAATAAGAGGGAGAGAAAA 2640
QY 2641 TCAAGAGATTTGATCAAGACAAATTTGAAACCAAGTGTCTAAGGATTTCTTTTATGAAAAAT 2700
DB 2641 TCAAGAGATTTGATCAAGACAAATTTGAAACCAAGTGTCTAAGGATTTCTTTTATGAAAAAT 2700
QY 2701 AGCAGCAGCACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTTCCCTTCA 2760
DB 2701 AGCAGCAGCACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTTCCCTTCA 2760
QY 2761 ATACTTTAGTAAACAGGAGCAAGGGGACCTGAGGTCACCTTCCCAAGGGGTTTCAAGCT 2820
DB 2761 ATACTTTAGTAAACAGGAGCAAGGGGACCTGAGGTCACCTTCCCAAGGGGTTTCAAGCT 2820
QY 2821 TCCAGCCAGCATGTAAACAAAGAGAAAGACGATAGGAAAGAGAAAGACGAGCTGAG 2880
DB 2821 TCCAGCCAGCATGTAAACAAAGAGAAAGACGATAGGAAAGAGAAAGACGAGCTGAG 2880
QY 2881 CAAAGTTAGGAAATCAACATTGAAATCCCAATGCAAGGAGTTCAACCCAGTTCCTTCTCT 2940

|||||
2881 CAGGTAGGAATCAACATTTGATCCCAATGAAAGGAGTTCAACCCACGTTCTTCTCT 2940
QY
2941 CAGCAAAAGCCTTCTATACCCCAACTTCACCTCGGCTCAAGCACAACCTTAGCCCATCT 3000
Db
2941 CAGCCAAAGCCTTCTATACCCCAACTTCACCTCGGCTCAAGCACAACCTTAGCCCATCT 3000
QY
3001 ATGGTGGTCAATCAACAGCCAACTTCAGTTTATATCTAGCCTGTTTGTGGTGCACCAAT 3060
Db
3001 ATGGTGGTCAATCAACAGCCAACTTCAGTTTATATCTAGCCTGTTTGTGGTGCACCAAT 3060
QY
3061 ATGATGATCCAGTCCAGTCCAGCCAGCGGTGAACTTTATACCCCAATACCTATGACG 3120
Db
3061 ATGATGATCCAGTCCAGTCCAGCCAGCGGTGAACTTTATACCCCAATACCTATGACG 3120
QY
3121 CCCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCAATATATGCCCAACAGCGG 3180
Db
3121 CCCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCAATATATGCCCAACAGCGG 3180
QY
3181 CAAGACAGCATCATCAGAGTGCATATGACACCCAGCGTCAGCAGCGGGCCCAACCGATT 3240
Db
3181 CAAGACAGCATCATCAGAGTGCATATGACACCCAGCGTCAGCAGCGGGCCCAACCGATT 3240
QY
3241 GCAGCCACCCACAGCTTACTCCACGCAATATGTTGCTTACAGTCCCTCAGCAGTCCCA 3300
Db
3241 GCAGCCACCCACAGCTTACTCCACGCAATATGTTGCTTACAGTCCCTCAGCAGTCCCA 3300
QY
3301 AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATAGT 3360
Db
3301 AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATAGT 3360
QY
3361 CCTGTATACAGGTATGCTAGAAATGATGCAACCAACACAGCCCGCCTGGTTTA 3420
Db
3361 CCTGTATACAGGTATGCTAGAAATGATGCAACCAACACAGCCCGCCTGGTTTA 3420
QY
3421 GTATCTTCTTCAGCAACTCAGTAGGGGCTCATGAGCAGAGCGATGCGATGATGCAATG 3480
Db
3421 GTATCTTCTTCAGCAACTCAGTAGGGGCTCATGAGCAGAGCGATGCGATGATGCAATG 3480
QY
3481 CCCAAATTAACATACAAACAGGAGACAGCCCTTCTTCTACTTTGCCATTTCCACGGGC 3540
Db
3481 CCCAAATTAACATACAAACAGGAGACAGCCCTTCTTCTACTTTGCCATTTCCACGGGC 3540
QY
3541 TCCCTTGTCTAGCAGTATGCGCACCTTAACGCTACCTTGACCCACATATCTCCACCCCT 3600
Db
3541 TCCCTTGTCTAGCAGTATGCGCACCTTAACGCTACCTTGACCCACATATCTCCACCCCT 3600
QY
3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCCCAACATGTTGGAGTCATCTGCA 3660
Db
3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCCCAACATGTTGGAGTCATCTGCA 3660
QY
3661 CCCAGTCTCTTACGACCATCAGCACCCAGCGCGCCAGGCTCTCCATCTGGCCAGTCCA 3720
Db
3661 CCCAGTCTCTTACGACCATCAGCACCCAGCGCGCCAGGCTCTCCATCTGGCCAGTCCA 3720
QY
3721 CAGCAGCAGTCAAGCAATTTACACCGGGGCTTGCGCAACTCCCACTCCATGACACCT 3780
Db
3721 CAGCAGCAGTCAAGCAATTTACACCGGGGCTTGCGCAACTCCCACTCCATGACACCT 3780
QY
3781 GCCTCCACACGCGAGTCCCAAGATAGTTTCCAGCAGCAGCAGCAGTCTCTTTACG 3840
Db
3781 GCCTCCACACGCGAGTCCCAAGATAGTTTCCAGCAGCAGCAGCAGTCTCTTTACG 3840
QY
3841 ATCCATCTCTCAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCCACTACT 3900
Db
3841 ATCCATCTCTCAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCCACTACT 3900
QY
3901 CAGGCTCATGTACAGTCAAGGAATGTTCTCTCATCCAACTGCGCCCAATGATG 3960
Db
3901 CAGGCTCATGTACAGTCAAGGAATGTTCTCTCATCCAACTGCGCCCAATGATG 3960
QY
3961 CTAATGAGCAGCAGCCACCGGGGTCGCCAGGCGCCCTCGCTCAAAGTGCACTACAG 4020
|||||

Db 3961 CTATAGCAGCAGACCCGCGGTCCTCCAGCGCGCTCGCTCAAAGTGCACACAG 4020
QY 4021 CCCATTCAGTCTCGACACAGCGCATTTCCCTTATATGACGCAACCTTCAGTACAAGCC 4080
Db 4021 CCCATTCAGTCTCGACACAGCGCATTTCCCTTATATGACGCAACCTTCAGTACAAGCC 4080
QY 4081 CACCACCAACAGCAGTTGTAAAGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCTC 4140
Db 4081 CACCACCAACAGCAGTTGTAAAGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCTC 4140
QY 4141 CTTTCTACTGCTTCTACCACTGGAAGCAGACAGAAAACCTAGAAATTTTATTTTGT 4200
Db 4141 CTTTCTACTGCTTCTACCACTGGAAGCAGACAGAAAACCTAGAAATTTTATTTTGT 4200
QY 4201 TAAATATATATGTTGATTTCTTCTTAACATCAATAGGAATGCTAAACAGTTTCACTTGCAG 4260
Db 4201 TAAATATATATGTTGATTTCTTCTTAACATCAATAGGAATGCTAAACAGTTTCACTTGCAG 4260
QY 4261 TGGAGATACTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTCCATATTTCCA 4320
Db 4261 TGGAGATACTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTCCATATTTCCA 4320
QY 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTCCGAAACTGGAAGTTATTT 4380
Db 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTCCGAAACTGGAAGTTATTT 4380
QY 4381 ATTTTAAATTAACCTTTGAAAGTCATGAACATCAACACATCAGCTAGCAAAAGTAACAAGT 4440
Db 4381 ATTTTAAATTAACCTTTGAAAGTCATGAACATCAACACATCAGCTAGCAAAAGTAACAAGT 4440
QY 4441 GATTTCTGCTATTACTGCTAAAAA 4481
Db 4441 GATTTCTGCTATTACTGCTAAAAA 4481

RESULT 2

AAZ23428
ID AAZ23428 standard; DNA; 4481 BP.

XX AAZ23428;

XX 19-JAN-2000 (first entry)

XX Human SCA2 DNA.

OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 163..4101
FT /*tag= a
FT /product= "SCA2"
XX WO9945944-A1.
XX 16-SEP-1999.
XX 11-MAR-1999; 99WO-US005250.
XX 12-MAR-1998; 98US-00041886.
XX (BURN-) BURNHAM INST.
XX Bredesen DE, Rabizadeh S;
XX WPI; 1999-561617/47.
XX P-FSDB; AAY3495.

XX PT New proapoptotic dependence peptides, used to develop products for
XX treating, e.g. Alzheimer's disease.
XX PS Disclosure; Page 130-135; 1999p; English.
XX CC This invention describes novel pure proapoptotic dependence peptides
CC which comprise a sequence of an active dependence domain selected from
CC dependence polypeptides consisting of p53NTR, androgen receptor, DCC,
CC huntingtin polypeptide, Machado-Joseph disease gene product, SCA1, SCA2,
CC SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable of
CC inducing cell death and can be used to develop products to mediate or
CC inhibit apoptosis. The methods can be used for reducing the severity of a
CC proapoptotic dependence domain mediated pathological conditions e.g.
CC Huntington's disease, Alzheimer's disease, Kennedy's disease,
CC Spinocerebellar ataxias, dentatorubropallidoluysian atrophy, Machado-
CC Joseph disease, stroke or head trauma. They can also be used for reducing
CC the severity of a pathological condition mediated by upregulated cell
CC proliferation or cell survival e.g. neoplastic, malignant, autoimmune or
CC fibrotic conditions. This sequence encodes the human SCA2 polypeptide
CC described in the method of the invention
XX
SQ Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
Query Match 100.0%; Score 4481; DB 2; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCCCGAGAAAGCAACCCAGCGCGCCGCTCTCTACAGTGTCCCTCCCGCCCGGG 60
DB 1 ACCCCGAGAAAGCAACCCAGCGCGCCGCTCTCTACAGTGTCCCTCCCGCCCGGG 60
QY 61 GCCACCTCAGTTCGCTTCGCTGCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
DB 61 GCCACCTCAGTTCGCTTCGCTGCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
QY 121 CACCTCCGCTCCACCCCGCGGCTCGCGCGCCGCGCCCTCCGATGGCTCAGCGCGCA 180
DB 121 CACCTCCGCTCCACCCCGCGGCTCGCGCGCCGCGCCCTCCGATGGCTCAGCGCGCGCA 180
QY 181 GCTCTCGAGTTCGCGGTGGCCACCGAGTCTCGCGCTTCGCCGACGAGTGGCC 240
DB 181 GCTCTCGAGTTCGCGGTGGCCACCGAGTCTCGCGCTTCGCCGACGAGTGGCC 240
QY 241 GGGTGGCGTCTCGCTCCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCG 300
DB 241 GGGTGGCGTCTCGCTCCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCG 300
QY 301 CCGGACCGGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 CCGGACCGGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CAGAGCTCCGCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 CAGAGCTCCGCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 TTTTCG 480
DB 421 TTTTCG 480
QY 481 GTCTTTCTCCCTCCG 540
DB 481 GTCTTTCTCCCTCCG 540
QY 541 CTGCG 600
DB 541 CTGCG 600
QY 601 TGTCCCG 660
DB 601 TGTCCCG 660
QY 661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720

DB 661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
QY 721 CAGCG 780
DB 721 CAGCG 780
QY 781 CCG 840
DB 781 CCG 840
QY 841 TCCTCGGTGCTCG 900
DB 841 TCCTCGGTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 AGTAAACAAAGGACTCGCTCAGCTACAGATTTCTTTGATGGAATCTATGCAAAATGAGG 960
DB 901 AGTAAACAAAGGACTCGCTCAGCTACAGATTTCTTTGATGGAATCTATGCAAAATGAGG 960
QY 961 ATGGTTCAATATCTTACATCAGTTGCTGCTCCAAATGTGAAGTACAAATGGA 1020
DB 961 ATGGTTCAATATCTTACATCAGTTGCTGCTCCAAATGTGAAGTACAAATGGA 1020
QY 1021 GGTATATATGAAGGAGTTTAAACCTTACAGTCCGAAAGTGTGATTTGGTACTTCATGCC 1080
DB 1021 GGTATATATGAAGGAGTTTAAACCTTACAGTCCGAAAGTGTGATTTGGTACTTCATGCC 1080
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTCGAGTTCGAGTAAAGTAAAGTAAAGT 1140
DB 1081 GCACATGAGAAAGTACAGAAATCCAGTTCGAGTTCGAGTAAAGTAAAGTAAAGT 1140
QY 1141 ATTTTGTTCAAATGTTTACAGTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
DB 1141 ATTTTGTTCAAATGTTTACAGTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1200
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 GAGAAAGGACCTGGAGCCCTGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1320
DB 1261 GAGAAAGGACCTGGAGCCCTGGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1320
QY 1321 TTGGAAGATGAGCTATCTAAATGGATGGATCCCAATGATATGTTTCGATATATGAGAA 1380
DB 1321 TTGGAAGATGAGCTATCTAAATGGATGGATCCCAATGATATGTTTCGATATATGAGAA 1380
QY 1381 AATTATGTTGTTAGTGTCTACAGTATGATAGCAGTTCCTTCGTATACAGTGCCTTTAGAA 1440
DB 1381 AATTATGTTGTTAGTGTCTACAGTATGATAGCAGTTCCTTCGTATACAGTGCCTTTAGAA 1440
QY 1441 AGAGATAACTCAGAGAAGATTTTAAACCGGGAAGCAAGGGCAACCCAGTTAGCAGAAGAA 1500
DB 1441 AGAGATAACTCAGAGAAGATTTTAAACCGGGAAGCAAGGGCAACCCAGTTAGCAGAAGAA 1500
QY 1501 ATTGAGTCAAGTGCAGCTCAAGCTCGAGTGGCGCTCGGAAATGATAGTAGGAGTGA 1560
DB 1501 ATTGAGTCAAGTGCAGCTCAAGCTCGAGTGGCGCTCGGAAATGATAGTAGGAGTGA 1560
QY 1561 GAAGAAAAATACACAGCAGTTCAGAGAAATTCAGTGAACGTTGAGGGGCAACAGCATAAAC 1620
DB 1561 GAAGAAAAATACACAGCAGTTCAGAGAAATTCAGTGAACGTTGAGGGGCAACAGCATAAAC 1620
QY 1621 ACTAGGAAAAATATAATATCTCTCGGACAAAGAAATAGAGAAAGTATATCTCTGGGA 1680
DB 1621 ACTAGGAAAAATATAATATCTCTCGGACAAAGAAATAGAGAAAGTATATCTCTGGGA 1680
QY 1681 AGTGGAGAGACAGAAATTCACCGGTATGGCGCAGCTGGATCGGGCTCCATGCCATCAAG 1740
DB 1681 AGTGGAGAGACAGAAATTCACCGGTATGGCGCAGCTGGATCGGGCTCCATGCCATCAAG 1740
QY 1741 TCCACTTCTCAGATTTTCAACCCGAAATTTCTGGTTTACAGCCAAAGAGTAGTTAAT 1800

QY 3961 CTAATGACGACACGACCGCCGGTCCCGAGCGCGCTCGCTCAAAGTGCACTACAG 4020
DB |||||||
QY 3961 CTAATGACGACACGACCGCCGGTCCCGAGCGCGCTCGCTCAAAGTGCACTACAG 4020
DB |||||||
QY 4021 CCATTTCAGTCTCGACCAACAGCGCATTTCCCTCTATATGACGACACCTTCAGTCAAGCC 4080
DB |||||||
QY 4021 CCATTTCAGTCTCGACCAACAGCGCATTTCCCTCTATATGACGACACCTTCAGTCAAGCC 4080
DB |||||||
QY 4081 CACCAACCAACAGCGTGTGAAGCTGCGCTGGAGAACCGAAAGCCAAATTCCTCTC 4140
DB |||||||
QY 4081 CACCAACCAACAGCGTGTGAAGCTGCGCTGGAGAACCGAAAGCCAAATTCCTCTC 4140
DB |||||||
QY 4141 CTTTCTACTGCTTCTACCACTGGAAGCAGACAGAACTAGAAATTCATTTATTTGTTTT 4200
DB |||||||
QY 4141 CTTTCTACTGCTTCTACCACTGGAAGCAGACAGAACTAGAAATTCATTTATTTGTTTT 4200
DB |||||||
QY 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTCACTTGCAG 4260
DB |||||||
QY 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTCACTTGCAG 4260
DB |||||||
QY 4261 TGAAGATATCTTGGACCGAGTAGAGCATTTAGGAACCTTGGGGCTATTCCATAATTCCA 4320
DB |||||||
QY 4261 TGAAGATATCTTGGACCGAGTAGAGCATTTAGGAACCTTGGGGCTATTCCATAATTCCA 4320
DB |||||||
QY 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380
DB |||||||
QY 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380
DB |||||||
QY 4381 ATTTTATATACCTTGAAGTCATGACACATCAGCTAGCAAAAGAGTAAACAGAGT 4440
DB |||||||
QY 4381 ATTTTATATACCTTGAAGTCATGACACATCAGCTAGCAAAAGAGTAAACAGAGT 4440
DB |||||||
QY 4441 GATTTCTGCTGCTATTACTGCTAAAAAANAAAAA 4481
DB |||||||
QY 4441 GATTTCTGCTGCTATTACTGCTAAAAAANAAAAA 4481
DB |||||||

RESULT 3
ADD18753
ID ADD18753 standard; DNA; 4481 BP.
XX
AC ADD18753;
XX
XX
DT 15-JAN-2004 (first entry)
DE Human disease related protein DNA sequence SeqID184.
XX human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnery; gene therapy;
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing; gene; ds.
XX
OS Homo sapiens.
XX
XX WO2003018621-A2.
XX
XX 06-MAR-2003.
XX
XX 23-AUG-2002; 2002WO-GB003892.
XX
XX 23-AUG-2001; 2001GB-00020558.
XX
XX 05-OCT-2001; 2001GB-00024037.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Kingeman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
XX
XX WPI; 2003-290046/28.
XX
XX P-PSDB; ADD18752.
XX

XX
PT New substantially purified polypeptide, useful for diagnosing or treating
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT wound healing.
XX
PS Claim 27; SEQ ID NO 184; 424pp; English.
XX
CC This invention relates to novel human genes and gene product which are
CC implicated in certain disease states. Compounds which modulate the
CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnerary activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transporation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
Query Match 100.0%; Score 4481; DB 10; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCCCCGAAGAAGCAACCCAGCGCGCGCTCTCACGTGTCCCTCCCGCCCCCGGG 60
DB |||||||
QY 1 ACCCCCGAAGAAGCAACCCAGCGCGCGCTCTCACGTGTCCCTCCCGCCCCCGGG 60
DB |||||||
QY 61 GCACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120
DB |||||||
QY 61 GCACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120
DB |||||||
QY 121 CACCTCCGCTCCACCCCGCGCTCGGCGCGCGCTCCGATGCGCTCAGCGCGCGCA 180
DB |||||||
QY 121 CACCTCCGCTCCACCCCGCGCTCGGCGCGCGCTCCGATGCGCTCAGCGCGCGCA 180
DB |||||||
QY 181 GCTCTCCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGCCC 240
DB |||||||
QY 181 GCTCTCCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGGTGGCCC 240
DB |||||||
QY 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCC 300
DB |||||||
QY 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCC 300
DB |||||||
QY 301 CCGGACCGTATCCCTCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB |||||||
QY 301 CCGGACCGTATCCCTCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB |||||||
QY 361 CAGAGTCCGCTCCCTCCGCTCAGACTGTTTGTAGCAACGCGCAACGCGCGCGCGCG 420
DB |||||||
QY 361 CAGAGTCCGCTCCCTCCGCTCAGACTGTTTGTAGCAACGCGCAACGCGCGCGCGCG 420
DB |||||||
QY 421 TTTTGGCGCGCGCTCCCGCGCGCTCTTGTCTCGCGCGCGCTCCCGCGCGCGCTTCGTCGTC 480
DB |||||||
QY 421 TTTTGGCGCGCGCTCCCGCGCGCTCTTGTCTCGCGCGCGCTCCCGCGCGCGCTTCGTCGTC 480
DB |||||||
QY 481 GTCTTCTCCCTCGCCAGCGCGCGCTCCCGCGCGCGCTCCCGCGCGCGCTCCCGCG 540
DB |||||||
QY 481 GTCTTCTCCCTCGCCAGCGCGCGCTCCCGCGCGCGCTCCCGCGCGCGCTCCCGCG 540
DB |||||||
QY 541 CTCGGCGCGCTCGCTCCCGCGCGCTTCGGGGTCTCTTGGCGCGCGCGCGCTCCCGCG 600
DB |||||||
QY 541 CTCGGCGCGCTCGCTCCCGCGCGCTTCGGGGTCTCTTGGCGCGCGCGCGCTCCCGCG 600
DB |||||||
QY 601 TGTTCCTCCCGCGCGCTCGCGCGCTGATGGCGCGCTTCCACATGCTCGTGAAGCGCGCG 660
DB |||||||
QY 601 TGTTCCTCCCGCGCGCTCGCGCGCTGATGGCGCGCTTCCACATGCTCGTGAAGCGCGCG 660
DB |||||||

Db 2821 TCAGGCCAGCATGTAAACAAGAGAGACGATAGGAAGAGAGAAAGACCCAGCTGAG 2880
Qy 2881 CAAGTTAGGAATCAACATTGTAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTTCT 2940
Db 2881 CAAGTTAGGAATCAACATTGTAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTTCT 2940
Qy 2941 CAGCCAAAGCCTTCTACTACCCCAACTTCACTTCGCGCTCAAGCACAACCTAGCCCATCT 3000
Db 2941 CAGCCAAAGCCTTCTACTACCCCAACTTCACTTCGCGCTCAAGCACAACCTAGCCCATCT 3000
Qy 3001 ATGGTGGGTCAACAACAGCAAACTCCAGTTTATCTCAGCCTGTGTTGTTTTCACCAAAAT 3060
Db 3001 ATGGTGGGTCAACAACAGCAAACTCCAGTTTATCTCAGCCTGTGTTGTTTTCACCAAAAT 3060
Qy 3061 ATGATGTATCCAGTCCAGTGCCGAGCCGCGTGCACACTTTTATACCCAATACCTATGAGC 3120
Db 3061 ATGATGTATCCAGTCCAGTGCCGAGCCGCGTGCACACTTTTATACCCAATACCTATGAGC 3120
Qy 3121 CCATGCCAGTGAATCAAGCCAGGAGATATAGAGCAGTACCAAAATATGCCCAAGCGG 3180
Db 3121 CCATGCCAGTGAATCAAGCCAGGAGATATAGAGCAGTACCAAAATATGCCCAAGCGG 3180
Qy 3181 CAAGCAGCAGCATCATCAGAGTGCATGATGACCCAGCGTCAAGCAGCGGCCACCGATT 3240
Db 3181 CAAGCAGCAGCATCATCAGAGTGCATGATGACCCAGCGTCAAGCAGCGGCCACCGATT 3240
Qy 3241 GCAGCCACCCCAACAGCTTACTCCAGCAATATGTTGCGCTACAGTCTCAGCAGTTCCCA 3300
Db 3241 GCAGCCACCCCAACAGCTTACTCCAGCAATATGTTGCGCTACAGTCTCAGCAGTTCCCA 3300
Qy 3301 AATCAGCCCTTGTTCAGATGTGCCACATTATCAGTCTCAGCATCTCATGCTATAGT 3360
Db 3301 AATCAGCCCTTGTTCAGATGTGCCACATTATCAGTCTCAGCATCTCATGCTATAGT 3360
Qy 3361 CTGTAAATACAGGTAATGCTAGAAATGAGTGGACACCAACACAGCCCGCTGTTTGA 3420
Db 3361 CTGTAAATACAGGTAATGCTAGAAATGAGTGGACACCAACACAGCCCGCTGTTTGA 3420
Qy 3421 GTATCTTCTTACAGCACTCAGTACGGGGCTCATGAGCAGACGCAATGCGATGTATGATG 3480
Db 3421 GTATCTTCTTACAGCACTCAGTACGGGGCTCATGAGCAGACGCAATGCGATGTATGATG 3480
Qy 3481 CCCAAATATACCAACAAGGAGACAAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3540
Db 3481 CCCAAATATACCAACAAGGAGACAAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3540
Qy 3541 TCCCTTGTCTCAGCAGTATGCGACCCCTAACGCTACCTGACCCACATACCTCCACACCT 3600
Db 3541 TCCCTTGTCTCAGCAGTATGCGACCCCTAACGCTACCTGACCCACATACCTCCACACCT 3600
Qy 3601 CAGCCTTTCAGCTACCCCTCAGTACGAGCAGCAAGCCAAATGCTGGAAGTCACTCTGCA 3660
Db 3601 CAGCCTTTCAGCTACCCCTCAGTACGAGCAGCAAGCCAAATGCTGGAAGTCACTCTGCA 3660
Qy 3661 CCAGTCTCTGTTACGACCATGAGCAACAGGCGCGCTCTCAGCTCTCAGCTGCGCAGTCA 3720
Db 3661 CCAGTCTCTGTTACGACCATGAGCAACAGGCGCGCTCTCAGCTCTCAGCTGCGCAGTCA 3720
Qy 3721 CAGCAGCAGTACGCAATTTACCAAGCGGGCTTGGCGCACTCCACCTCCATGACACCT 3780
Db 3721 CAGCAGCAGTACGCAATTTACCAAGCGGGCTTGGCGCACTCCACCTCCATGACACCT 3780
Qy 3781 GCCTCCAAACAGCAGTCCGCAAGAAATAGTTTCCAGCAGCACAACAGACTGTCTTTACG 3840
Db 3781 GCCTCCAAACAGCAGTCCGCAAGAAATAGTTTCCAGCAGCACAACAGACTGTCTTTACG 3840
Qy 3841 ATCCATCTCTTCTCAGCTTACGCGGGGTATACCAACCCACCCCAATGCGCCACGTAACCT 3900
Db 3841 ATCCATCTCTTCTCAGCTTACGCGGGGTATACCAACCCACCCCAATGCGCCACGTAACCT 3900
Qy 3901 CAGGCTCATGTACAGTCAGGAATGCTTCTTCTCATCCCACTGCCCATGCGCCAATGATG 3960
Db 3901 CAGGCTCATGTACAGTCAGGAATGCTTCTTCTCATCCCACTGCCCATGCGCCAATGATG 3960

Qy 3961 CTAATGACGACACAGCCACCCGCGGTCCCHAGGCGCGCTCGCTCAAAGTGCACCTACAG 4020
Db 3961 CTAATGACGACACAGCCACCCGCGGTCCCHAGGCGCGCTCGCTCAAAGTGCACCTACAG 4020
Qy 4021 CCCATTCCAGTCTCCAGCAACAGCGCATTTCCCTTATATGACGACACCTTTCAGTACAAGCC 4080
Db 4021 CCCATTCCAGTCTCCAGCAACAGCGCATTTCCCTTATATGACGACACCTTTCAGTACAAGCC 4080
Qy 4081 CACCACCAACAGCAGTGTGTAAGGCTGCCCTGGAGAACCGAAAGGCCAAATTCCTCTCTC 4140
Db 4081 CACCACCAACAGCAGTGTGTAAGGCTGCCCTGGAGAACCGAAAGGCCAAATTCCTCTCTC 4140
Qy 4141 CTTCTACTGCTTCTACCAACTGGAGACACAGAAACTAGAAATTTTATTTTCTTTT 4200
Db 4141 CTTCTACTGCTTCTACCAACTGGAGACACAGAAACTAGAAATTTTATTTTCTTTT 4200
Qy 4201 TAAAATATATATGTTGATTCTTTGTAACATCAATAGGAATGCTAACAGTTCACATTGCGAG 4260
Db 4201 TAAAATATATATGTTGATTCTTTGTAACATCAATAGGAATGCTAACAGTTCACATTGCGAG 4260
Qy 4261 TGGAAAGATACCTTGGACCGAGTAGAGGCATTTTAGGAACTTTGGGGGCTATTTCCTAATAATCCA 4320
Db 4261 TGGAAAGATACCTTGGACCGAGTAGAGGCATTTTAGGAACTTTGGGGGCTATTTCCTAATAATCCA 4320
Qy 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380
Db 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGCAGAACTGGAAGTTATTT 4380
Qy 4381 ATTTTAAATACCTTGAAGTCAATGACACATCAGCTAGCAAAAGAGTAACAAGAGT 4440
Db 4381 ATTTTAAATACCTTGAAGTCAATGACACATCAGCTAGCAAAAGAGTAACAAGAGT 4440
Qy 4441 GATTCCTGCTGCTATTACTGCTTAAATAAAAAAAAAAAAAAAAAAAAA 4481
Db 4441 GATTCCTGCTGCTATTACTGCTTAAATAAAAAAAAAAAAAAAAAAAAA 4481

RESULT 4

ADH69466

ID ADH69466 standard; DNA; 4481 BP.

XX ADH69466;

XX DT 25-MAR-2004 (first entry)

XX Human spinocerebellar ataxia 2; SCA2; gene.

XX Spinocerebellar ataxia 2; SCA2; genetic analysis; gene; ds; human.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 163..4101

XX /tag= a

XX /product= "Human SCA2 protein"

XX US6623927-B1.

XX 23-SEP-2003.

XX 08-NOV-2000; 2000US-00707919.

XX 08-NOV-2000; 2000US-00707919.

XX (COUL) COUNCIL SCI & IND RES.

XX Brahmachari SK, Choudhry S, Mukerji M, Jain S;

XX WPI; 2003-895465/82.

XX P-PSDB; ADH69465.

XX New oligonucleotide primers useful in predicting an individual's

PT susceptibility to spinocerebellar ataxia 2 (SCA2) disease, in genetic
PT analysis of SCA2 gene in a population, and in detecting allelic variants
PT of the SCA2 gene.

PS Disclosure; SEQ ID NO 20; 30pp; English.

XX The invention relates to an oligonucleotide primer. The oligonucleotide
CC primers are useful in molecular diagnosis, in predicting an individual's
CC susceptibility to spinocerebellar ataxia 2 (SCA2) disease, in genetic
CC analysis of SCA2 gene in a population and in detecting allelic variants
CC of the SCA2 gene. The present sequence is human SCA2 gene.

XX Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

Query Match 100.0%; Score 4481; DB 10; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTTGTCCTTCCGCGCCCGGG	60
DB	1	ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTTGTCCTTCCGCGCCCGGG	60
QY	61	GCACCTCAGGTTCTGCTTCCGTCGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG	120
DB	61	GCACCTCAGGTTCTGCTTCCGTCGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG	120
QY	121	CACCTCCGCTCCACCCGCGCTCGGCGCGCCCGCTCCGATCGCGTCCGATCGCGCCGCA	180
DB	121	CACCTCCGCTCCACCCGCGCTCGGCGCGCCCGCTCCGATCGCGTCCGATCGCGCCGCA	180
QY	181	GCTCTCGAGTCCCGGTTGGCCACCGAGTCTCGCGCTTCGCGACGACGAGTGGGCC	240
DB	181	GCTCTCGAGTCCCGGTTGGCCACCGAGTCTCGCGCTTCGCGACGACGAGTGGGCC	240
QY	241	GSGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCG	300
DB	241	GSGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCG	300
QY	301	CGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG	360
DB	301	CGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG	360
QY	361	CAGACTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACCGGCGGCGGCGGCGG	420
DB	361	CAGACTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACCGGCGGCGGCGGCGG	420
QY	421	TTTGGCGCGGCTCCGCGCGCTCTGCTCGGCGGCGCTCCCGCGCGCGCTTGGTCTC	480
DB	421	TTTGGCGCGGCTCCGCGCGCTCTGCTCGGCGGCGCTCCCGCGCGCGCTTGGTCTC	480
QY	481	GTCCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	540
DB	481	GTCCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	540
QY	541	CTGCGCGCGGTCGTCGCGCGGTCGCGGCTCTTGGCGCGCGCGCGCGCGCGCG	600
DB	541	CTGCGCGCGGTCGTCGCGCGGTCGCGGCTCTTGGCGCGCGCGCGCGCGCGCG	600
QY	601	TGTCGCCCGCGGTCGAGCGGTTATGSGGCCCTCACCATGTCGCTCAAGCGCCG	660
DB	601	TGTCGCCCGCGGTCGAGCGGTTATGSGGCCCTCACCATGTCGCTCAAGCGCCG	660
QY	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720
DB	661	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	720
QY	721	CAGCGCGCGCGGCTCGCAATGTCGCAAGCCCGCGCGGCGCGCTTCTAGCGTCG	780
DB	721	CAGCGCGCGCGGCTCGCAATGTCGCAAGCCCGCGCGGCGCGCTTCTAGCGTCG	780
QY	781	CCGCGCGCGCGCTTCCGCGCTCTGCTCGGCTCTCTGCTCTCGGCGCGCGCTCCC	840
DB	781	CCGCGCGCGCGCTTCCGCGCTCTGCTCGGCTCTCTGCTCTCGGCGCGCGCTCCC	840

QY	841	TCCTCGTGGTGGCGGCGACCTTCGGCGCGGGAGGCCCGGCTTCGGGAGAGGTCCAAAC	900
DB	841	TCCTCGTGGTGGCGGCGACCTTCGGCGCGGGAGGCCCGGCTTCGGGAGAGGTCCAAAC	900
QY	901	AGTAACAAAGGAGTGCCTCAGTCTACGATTTCTTTTGTGATGAATCTATGCAAAATGAGG	960
DB	901	AGTAACAAAGGAGTGCCTCAGTCTACGATTTCTTTTGTGATGAATCTATGCAAAATGAGG	960
QY	961	ATGTTTCATATACATACATCAGTTGTTGGCTCCAAATCTGAAGTCAAGTGAAGAAATCGA	1020
DB	961	ATGTTTCATATACATACATCAGTTGTTGGCTCCAAATCTGAAGTCAAGTGAAGAAATCGA	1020
QY	1021	GCTATATGAAGAGTGTAAATACTTACAGTCCGAAGTGTGATTTGGTACTTCATGCC	1080
DB	1021	GCTATATGAAGAGTGTAAATACTTACAGTCCGAAGTGTGATTTGGTACTTCATGCC	1080
QY	1081	GCACATGAGAAAGTACAGAAATCCAGTTCGGGCGGAAACCTGAAGAAATATGAGAGT	1140
DB	1081	GCACATGAGAAAGTACAGAAATCCAGTTCGGGCGGAAACCTGAAGAAATATGAGAGT	1140
QY	1141	ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTACAGTTTAAAGATATGAGCTTAT	1200
DB	1141	ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTACAGTTTAAAGATATGAGCTTAT	1200
QY	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGAAGTGAATGGGCAACACAAA	1260
DB	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGAAGTGAATGGGCAACACAAA	1260
QY	1261	GAGAAAGACCTGGAGCCCTGGGATGCGAGTGAACTCACAGCAATGAGGAACTTGAGGCT	1320
DB	1261	GAGAAAGACCTGGAGCCCTGGGATGCGAGTGAACTCACAGCAATGAGGAACTTGAGGCT	1320
QY	1321	TTGAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATAATGAAGA	1380
DB	1321	TTGAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATAATGAAGA	1380
QY	1381	AATTATGTTAGTGTCTACGTATGATAGAGTGTATCTTCGTATACAGTGCCTTAGAA	1440
DB	1381	AATTATGTTAGTGTCTACGTATGATAGAGTGTATCTTCGTATACAGTGCCTTAGAA	1440
QY	1441	AGAGATAACTCAGAAAGAAATTTTAAACGGGAGCAAGGGCAACACAGTTAGCAAGAA	1500
DB	1441	AGAGATAACTCAGAAAGAAATTTTAAACGGGAGCAAGGGCAACACAGTTAGCAAGAA	1500
QY	1501	ATTGAGTCAAGTGCCTCAGTACAAAGCTCGAGTGCCTCGGAAATGATAGAGGTCAG	1560
DB	1501	ATTGAGTCAAGTGCCTCAGTACAAAGCTCGAGTGCCTCGGAAATGATAGAGGTCAG	1560
QY	1561	GAAGAAAATAACACAGCAGTTTCAGAGAAATTCAGTGAACCTGAGGGGACAGCATAAAC	1620
DB	1561	GAAGAAAATAACACAGCAGTTTCAGAGAAATTCAGTGAACCTGAGGGGACAGCATAAAC	1620
QY	1621	ACTAGGAAAATAAATATATTCCTCTGGA CAAGAAATAGAGAGTCAATTCCTGGGGA	1680
DB	1621	ACTAGGAAAATAAATATATTCCTCTGGA CAAGAAATAGAGAGTCAATTCCTGGGGA	1680
QY	1681	AGTGGAGACAGAAATTCACCGGTATGGCCAGCTGGATCGGCTCCATCCATCAAGA	1740
DB	1681	AGTGGAGACAGAAATTCACCGGTATGGCCAGCTGGATCGGCTCCATCCATCAAGA	1740
QY	1741	TCCACTTCTCACACTTCAGATTTTCAACCCGAAATTCGTTTCAGACCAAGAGTGTAAAT	1800
DB	1741	TCCACTTCTCACACTTCAGATTTTCAACCCGAAATTCGTTTCAGACCAAGAGTGTAAAT	1800
QY	1801	GGAGGTGTTCCCTGGCAATTCGCTTGGCCATCTCTCTCTCGGCCACCTTCTCGCTAC	1860
DB	1801	GGAGGTGTTCCCTGGCAATTCGCTTGGCCATCTCTCTCTCGGCCACCTTCTCGCTAC	1860
QY	1861	CAGTCAAGTCCCACTCTTCCACCTCGGCGACGCCCTACACGGCGCCCTCCAGG	1920
DB	1861	CAGTCAAGTCCCACTCTTCCACCTCGGCGACGCCCTACACGGCGCCCTCCAGG	1920

QY 1921 CCCCTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCTCATGGTTCTCCAGCTCT 1980
DB |||||
QY 1921 CCCCTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCTCATGGTTCTCCAGCTCT 1980
DB |||||
QY 1981 GTCTCTACTATGCTTAAAGCGCATGTCTTCAGAGAGGCGCTCCAGAGATGTCCCAAGGCC 2040
DB |||||
QY 1981 GTCTCTACTATGCTTAAAGCGCATGTCTTCAGAGAGGCGCTCCAGAGATGTCCCAAGGCC 2040
DB |||||
QY 2041 CAGCGACATCTTCGAAATCAGAGATTTCTGTGGAGGGGTTCANATCCAGTGGCCTA 2100
DB |||||
QY 2041 CAGCGACATCTTCGAAATCAGAGATTTCTGTGGAGGGGTTCANATCCAGTGGCCTA 2100
DB |||||
QY 2101 GAAATTTGATCCACACACCCAGTGAAGCAGTACTCTCCAGTGAAGGACCACT 2160
DB |||||
QY 2101 GAAATTTGATCCACACACCCAGTGAAGCAGTACTCTCCAGTGAAGGACCACT 2160
DB |||||
QY 2161 CCTCGGGGGGAAAGTGGTCTCATGAGTGGTTCAGTGGGGTCCAGATTAATCCCTTAAACT 2220
DB |||||
QY 2161 CCTCGGGGGGAAAGTGGTCTCATGAGTGGTTCAGTGGGGTCCAGATTAATCCCTTAAACT 2220
DB |||||
QY 2221 CATAGACCCAGGTCTCCAGACAGAAAGTATTTGAAATACCCCGAGTGGGCGAGTTCTT 2280
DB |||||
QY 2221 CATAGACCCAGGTCTCCAGACAGAAAGTATTTGAAATACCCCGAGTGGGCGAGTTCTT 2280
DB |||||
QY 2281 GCTTCTCCCAAGCTGGTATTAATCCAACTGAAGCTGTTGCCATGCTATTCAGCTGCA 2340
DB |||||
QY 2281 GCTTCTCCCAAGCTGGTATTAATCCAACTGAAGCTGTTGCCATGCTATTCAGCTGCA 2340
DB |||||
QY 2341 TCTCTCTACGCTGTCTGATCTGATCGAAGCAGAGCTGTTACCCCTCTTAGTGAGGCTAAA 2400
DB |||||
QY 2341 TCTCTCTACGCTGTCTGATCTGATCGAAGCAGAGCTGTTACCCCTCTTAGTGAGGCTAAA 2400
DB |||||
QY 2401 GATTCAGGCTTCAAGATCAGAGGCGAGAACTCTCTGCGAGGGAATAAGAAATATTTAAA 2460
DB |||||
QY 2401 GATTCAGGCTTCAAGATCAGAGGCGAGAACTCTCTGCGAGGGAATAAGAAATATTTAAA 2460
DB |||||
QY 2461 CCATATGAACATCACTAGTCTCTCAAGCTGAAAGCTGAAAGAGTATATCACAGTTGTT 2520
DB |||||
QY 2461 CCATATGAACATCACTAGTCTCTCAAGCTGAAAGCTGAAAGAGTATATCACAGTTGTT 2520
DB |||||
QY 2521 TCTGACATAGAAAACAGATGATGATTTAAAGAAATTTAAAGATGATTTAGTTACAG 2580
DB |||||
QY 2521 TCTGACATAGAAAACAGATGATGATTTAAAGAAATTTAAAGATGATTTAGTTACAG 2580
DB |||||
QY 2581 CCAAGTTCTACTTCTGAATCTATGATCAACTACTAAACAAAATAGAGGGAGAAAAA 2640
DB |||||
QY 2581 CCAAGTTCTACTTCTGAATCTATGATCAACTACTAAACAAAATAGAGGGAGAAAAA 2640
DB |||||
QY 2641 TCAAGAGATTTGATCAAGACAAAATTTGAACCAAGTCTAAGGATTTCTTTTCAATGAAAT 2700
DB |||||
QY 2641 TCAAGAGATTTGATCAAGACAAAATTTGAACCAAGTCTAAGGATTTCTTTTCAATGAAAT 2700
DB |||||
QY 2701 AGCAGCAGCACTGTACCGTGGCAGCAGCAAGCCGAATAGCCCGATTTCCCTTCA 2760
DB |||||
QY 2701 AGCAGCAGCACTGTACCGTGGCAGCAGCAAGCCGAATAGCCCGATTTCCCTTCA 2760
DB |||||
QY 2761 ATACTTAGTAAACCGGAGCAAAAGAGGGACCTGAGTCACTTCCCAAGGGGTTCAGACT 2820
DB |||||
QY 2761 ATACTTAGTAAACCGGAGCAAAAGAGGGACCTGAGTCACTTCCCAAGGGGTTCAGACT 2820
DB |||||
QY 2821 TCCAGGCCAGCATGTAAACAGAGAAAGAGATAGAGAGAGAAAGAGCGAGCTGAG 2880
DB |||||
QY 2821 TCCAGGCCAGCATGTAAACAGAGAAAGAGATAGAGAGAGAAAGAGCGAGCTGAG 2880
DB |||||
QY 2881 CAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGGAGTTCAACCCAGTCTCTCTCT 2940
DB |||||
QY 2881 CAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGGAGTTCAACCCAGTCTCTCTCT 2940
DB |||||
QY 2941 CAGCCAAAGCCTTCTACTAGCCCAACTTCACTCTGGGCTCAAGCACAACCTAGCCCATCT 3000
DB |||||
QY 2941 CAGCCAAAGCCTTCTACTAGCCCAACTTCACTCTGGGCTCAAGCACAACCTAGCCCATCT 3000
DB |||||
QY 3001 ATGGTGGGTCACTCAAGCCAACTCCAGTTTATCTCAGCCTGTGTTTGGCACCAAT 3060

DB |||||
QY 3001 ATGGTGGGTCACTCAAGCCAACTCCAGTTTATATCTCAGCCTGTTTGTTCACCAAAAT 3060
DB |||||
QY 3061 ATGATGTATCCAGTCCCAGTGAGCCAGCGGTGAAACCTTTATACCAATACCTATGAGC 3120
DB |||||
QY 3061 ATGATGTATCCAGTCCCAGTGAGCCAGCGGTGAAACCTTTATACCAATACCTATGAGC 3120
DB |||||
QY 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCCAACAGCG 3180
DB |||||
QY 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCCCCAACAGCG 3180
DB |||||
QY 3181 CAAAGACCCAGTCATCAGAGTGCCATGATGCAACCCAGCGTTCAGAGCGGGCCACCGATT 3240
DB |||||
QY 3181 CAAAGACCCAGTCATCAGAGTGCCATGATGCAACCCAGCGTTCAGAGCGGGCCACCGATT 3240
DB |||||
QY 3241 GCAGCCACCCAGTCTTACTCCAGCGAAATATGTTGCCCTACAGTCTCAGCAGTTCCCA 3300
DB |||||
QY 3241 GCAGCCACCCAGTCTTACTCCAGCGAAATATGTTGCCCTACAGTCTCAGCAGTTCCCA 3300
DB |||||
QY 3301 AATCAGCCCTTGTTCAGCATGTGCAATATCAGTCTCAGCATCCTCATGCTATAGT 3360
DB |||||
QY 3301 AATCAGCCCTTGTTCAGCATGTGCAATATCAGTCTCAGCATCCTCATGCTATAGT 3360
DB |||||
QY 3361 CCTGTAATACAGGGTAAATGCTAGAATGATGGCAACCAACACACGCCAGCCTGTTTTA 3420
DB |||||
QY 3361 CCTGTAATACAGGGTAAATGCTAGAATGATGGCAACCAACACACGCCAGCCTGTTTTA 3420
DB |||||
QY 3421 GTATCTTCTTACAGCAACTCAGTACGGGGTCTATGAGCAGCATGCGATGTATGATGT 3480
DB |||||
QY 3421 GTATCTTCTTACAGCAACTCAGTACGGGGTCTATGAGCAGCATGCGATGTATGATGT 3480
DB |||||
QY 3481 CCAAAATTTACCATCAAAAGGAGCAAGCCCTTCTTTTACTTTGCGCATTTCCACGGGC 3540
DB |||||
QY 3481 CCAAAATTTACCATCAAAAGGAGCAAGCCCTTCTTTTACTTTGCGCATTTCCACGGGC 3540
DB |||||
QY 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTTACCTGACCCACATACTCCACACCT 3600
DB |||||
QY 3541 TCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTTACCTGACCCACATACTCCACACCT 3600
DB |||||
QY 3601 CAGCCTTACAGTACCCCCCTGAGCAGCAGCAAGCAACATGTTGGAGTCACTCTGCA 3660
DB |||||
QY 3601 CAGCCTTACAGTACCCCCCTGAGCAGCAGCAAGCAACATGTTGGAGTCACTCTGCA 3660
DB |||||
QY 3661 CCGAGTCTCTTACGACCATCAGCAGCAGCGGGCCAGGCTCTCCATCTGGCAGTCCA 3720
DB |||||
QY 3661 CCGAGTCTCTTACGACCATCAGCAGCAGCGGGCCAGGCTCTCCATCTGGCAGTCCA 3720
DB |||||
QY 3721 CAGCAGCAGTCAAGCCTTTTACAGCGGGCTTGGCCCAACTCCACCTCCATGACACCT 3780
DB |||||
QY 3721 CAGCAGCAGTCAAGCCTTTTACAGCGGGCTTGGCCCAACTCCACCTCCATGACACCT 3780
DB |||||
QY 3781 GCTTCCAAACAGCAGTCCGACAGAAATAGTTTCCCAGCAGCACAACAGACTGTCTTACG 3840
DB |||||
QY 3781 GCTTCCAAACAGCAGTCCGACAGAAATAGTTTCCCAGCAGCACAACAGACTGTCTTACG 3840
DB |||||
QY 3841 ATCCATCTCTTACGCTTACGCGGGCGTATACCAACCCACCCACATGCGCCACAGTACCT 3900
DB |||||
QY 3841 ATCCATCTCTTACGCTTACGCGGGCGTATACCAACCCACCCACATGCGCCACAGTACCT 3900
DB |||||
QY 3901 CAGGCTCATGTACAGTCAGAAATGTTCTTCTCATCCAACTGCCCATCGCCAATGATG 3960
DB |||||
QY 3901 CAGGCTCATGTACAGTCAGAAATGTTCTTCTCATCCAACTGCCCATCGCCAATGATG 3960
DB |||||
QY 3961 CTAATGACACACAGCAGCAGCCCGGGTCCCGAGCGCGCTCTCGCTCAAAAGTGCACACAG 4020
DB |||||
QY 3961 CTAATGACACACAGCAGCAGCCCGGGTCCCGAGCGCGCTCTCGCTCAAAAGTGCACACAG 4020
DB |||||
QY 4021 CCAATTTCCAGTCTCGACAAAGCGCATTTCCCTTATATGAGCAGCCCTTTCAGTACAAGCC 4080
DB |||||
QY 4021 CCAATTTCCAGTCTCGACAAAGCGCATTTCCCTTATATGAGCAGCCCTTTCAGTACAAGCC 4080
DB |||||
QY 4081 CACCAACACAGCAGTGTGAAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCTCTC 4140

Db 4081 CACCACCAACAGCAGTGTGAAGCTGCCTGGAGAACCGAAAGGCCAAATTCCTCTCCTC 4140
Qy 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGACGAAACTAGAAATTCATTTATTTTGTTTT 4200
Db 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGACGAAACTAGAAATTCATTTATTTTGTTTT 4200
Qy 4201 TAAATATATATGTTGATTCTTGTAAACATCCAATAGGAATGCTTAACAGATTTCACATTGCAG 4260
Db 4201 TAAATATATATGTTGATTCTTGTAAACATCCAATAGGAATGCTTAACAGATTTCACATTGCAG 4260
Qy 4261 TGGAGATACATTGGACCGAGTAGAGCAATTTAGGAACTTTGGGGGCTATTCCATAATTCCTCA 4320
Db 4261 TGGAGATACATTGGACCGAGTAGAGCAATTTAGGAACTTTGGGGGCTATTCCATAATTCCTCA 4320
Qy 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGC CGAAACTGGAAAGTTATTT 4380
Db 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGC CGAAACTGGAAAGTTATTT 4380
Qy 4381 ATTTTAAATACCCCTTGAAAGTCATGAACATCAGCTAGCATAAGAAAGTAAACAAGAGT 4440
Db 4381 ATTTTAAATACCCCTTGAAAGTCATGAACATCAGCTAGCATAAGAAAGTAAACAAGAGT 4440
Qy 4441 GATTCCTGCTGCTATTACTTGCTATAAAAA 4481
Db 4441 GATTCCTGCTGCTATTACTTGCTATAAAAA 4481

RESULT 5
ABV75291
ID ABV75291 standard; DNA; 4481 BP.
XX
AC ABV75291;
XX
DT 07-MAR-2003 (first entry)
XX
DE Human SCA2 polypeptide encoding DNA.
XX
KW SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity;
KW memory impairment; mutant; human; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 163..4101
FT /*tag= a
FT /*product= "SCA2"
XX
PN WO200289567-A2.
XX
PD 14-NOV-2002.
XX
PF 07-MAY-2002; 2002WO-US014698.
XX
PR 07-MAY-2001; 2001US-0289231P.
XX
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI Pulst SM;
XX
XX WPI; 2003-103467/09.
DR P-P5DB; ABB82698.
XX
PT New mutant mouse, useful as a disease model for identifying agents for
PT treating obesity or memory impairment, comprises a disrupted
PT spinocerebellar ataxia 2 gene.
XX
PS Disclosure; Page 65-72; 76pp; English.
XX

The invention relates to a mutant non-human mammal having a disrupted spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a disease model for identifying agents or compounds that may treat obesity or memory impairment. The stem cell is used for generating the mutant animal. The present sequence represents the nucleotide sequence (GenBank

CC Accession No. U70323) encoding the human SCA2 polypeptide
XX
SQ Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
Query Match 100.0%; Score 4481; DB 10; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACCCCCAGAGAAAGCAACCCAGCGCGCCGCTCTCTCACTGCTTCCGCTAAAGAGTCCCTATCCG 120
Db 1 ACCCCCAGAGAAAGCAACCCAGCGCGCGCCGCTCTCTCACTGCTTCCGCTAAAGAGTCCCTATCCG 60
Qy 61 GCCACTCAGCTTCTGCTTCCGCTGACCCCTCCGACTTCCGCTAAAGAGTCCCTATCCG 120
Db 61 GCCACTCAGCTTCTGCTTCCGCTGACCCCTCCGACTTCCGCTAAAGAGTCCCTATCCG 120
Qy 121 CACCTCCGCTCCCAACCCGCGCCCTCGCGCGCCGCTCTCGATGCGCTCAGAGCGCCCA 180
Db 121 CACCTCCGCTCCCAACCCGCGCCCTCGCGCGCCGCTCTCGATGCGCTCAGAGCGCCCA 180
Qy 181 GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTTCGCGCAGCAGGTGGGCC 240
Db 181 GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCGCTTTCGCGCAGCAGGTGGGCC 240
Qy 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy 301 CCGGAGCGATATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 CCGGAGCGATATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAGCGCGCGCGCGCG 420
Db 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGGCAGCGCGCGCGCGCG 420
Qy 421 TTTTGGCGCGGCTCGCGCGGCTCTTGGTCTCGCGCGGCTCTCCCGCGCGCTTCTGTCGT 480
Db 421 TTTTGGCGCGGCTCTCGCGGCTCTTGGTCTCGCGCGGCTCTCCCGCGCGCTTCTGTCGT 480
Qy 481 GTCTTCTCCCTCCG 540
Db 481 GTCTTCTCCCTCCG 540
Qy 541 CTGCGCGCGCTGCTTCCCGCGCGCTTCCCGCGCGCTCTCTTGGCGCGCGCGCGCGCG 600
Db 541 CTGCGCGCGCTGCTTCCCGCGCGCTTCCCGCGCGCTCTCTTGGCGCGCGCGCGCGCG 600
Qy 601 TGTCCCGCGCGCTGCGAGCGGTGTATGGGCGCTTCCACCATGTCTGCTGAAGCGCCAG 660
Db 601 TGTCCCGCGCGCTGCGAGCGGTGTATGGGCGCTTCCACCATGTCTGCTGAAGCGCCAG 660
Qy 661 CAG 720
Db 661 CAG 720
Qy 721 CAGCAGCGCGCGCGCTGCGCAATGTCCGCAAGCGCGCGCGCGCGCGCGCTTCTAGCGT 780
Db 721 CAGCAGCGCGCGCGCGCTGCGCAATGTCCGCAAGCGCGCGCGCGCGCGCTTCTAGCGT 780
Qy 781 CCGCGCGCGCGCGCTTCCGCGCTCTGCTCTGCTCTCTGCTCTGCTCTGCTCTGCTCTCC 840
Db 781 CCGCGCGCGCGCGCTTCCGCGCTCTGCTCTGCTCTCTGCTCTGCTCTGCTCTGCTCTCC 840
Qy 841 TCCTCGGTGTCGCGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 841 TCCTCGGTGTCGCGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy 901 AGTAACAAAGGACTCGCTCAGTCTACGATTTCTTTTGTGATGAATCTATGCAAAATATGAGG 960
Db 901 AGTAACAAAGGACTCGCTCAGTCTACGATTTCTTTTGTGATGAATCTATGCAAAATATGAGG 960
Qy 961 ATGGTTTATATATCTTACATCATGTTGTTGGCTCCAAATGTGAAGTACAAGTAAAAAATGGA 1020

QY 3181 CAAGACCAAGCATCATCAGAGTGCATGATGCAACCCAGCGTCAGCAGCGGGCCACCGGATT 3240
Db 3181 CAAGACCAAGCATCATCAGAGTGCATGATGCAACCCAGCGTCAGCAGCGGGCCACCGGATT 3240
QY 3241 CGAGCCACCCACAGCTTACTCCACGCAATATGTTGCTCAGCTCCTCAGCAGTTCCCA 3300
Db 3241 CGAGCCACCCACAGCTTACTCCACGCAATATGTTGCTCAGCTCCTCAGCAGTTCCCA 3300
QY 3301 AATCAGCCCTTGTTCAGCATGTGCAATATATCAGTCTCAGCATCCTCATGTCTATAGT 3360
Db 3301 AATCAGCCCTTGTTCAGCATGTGCAATATATCAGTCTCAGCATCCTCATGTCTATAGT 3360
QY 3361 CCTGTAATACAGGGTAATGCTAGAAATGATGGCAACCAACCAACAGCAGCCGCTGGTTA 3420
Db 3361 CCTGTAATACAGGGTAATGCTAGAAATGATGGCAACCAACCAACAGCAGCCGCTGGTTA 3420
QY 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCATGCGATGTATGATGT 3480
Db 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCATGCGATGTATGATGT 3480
QY 3481 CCCAAATTTACCATCAACAAGGAGACAAGCCCTTCTTCTACTTTGCCATTTTCCACGGGC 3540
Db 3481 CCCAAATTTACCATCAACAAGGAGACAAGCCCTTCTTCTACTTTGCCATTTTCCACGGGC 3540
QY 3541 TCCCTTGTCTCAGCAGTATGCGACCCCTAACCGCTACCTTGCACCCACATACTCCACCCCT 3600
Db 3541 TCCCTTGTCTCAGCAGTATGCGACCCCTAACCGCTACCTTGCACCCACATACTCCACCCCT 3600
QY 3601 CAGCCTTCAGCTTACCCCACTGGACAGCAGCAAGCCCAACATGTTGGAAGTCATCTCTCA 3660
Db 3601 CAGCCTTCAGCTTACCCCACTGGACAGCAGCAAGCCCAACATGTTGGAAGTCATCTCTCA 3660
QY 3661 CCCAGTCTCTTTCAGCACCATCAGCACCAAGCCCGCCAGGCTCTCCATCTGCGCAGTCCA 3720
Db 3661 CCCAGTCTCTTTCAGCACCATCAGCACCAAGCCCGCCAGGCTCTCCATCTGCGCAGTCCA 3720
QY 3721 CAGCAGCAGTCAAGCATTTTCAACCGGGGTTGCGCAACTCCACCTCCATGACACCT 3780
Db 3721 CAGCAGCAGTCAAGCATTTTCAACCGGGGTTGCGCAACTCCACCTCCATGACACCT 3780
QY 3781 GCCTCCACACGCGTCCGACAGATAGTTTCCAGCAGCAGCAGCAGTCTCTTTAGC 3840
Db 3781 GCCTCCACACGCGTCCGACAGATAGTTTCCAGCAGCAGCAGCAGTCTCTTTAGC 3840
QY 3841 ATCCATCTCTTCAGCTTCAGCGCGGTATACCAACCCACCCCATGCGCCACGTACT 3900
Db 3841 ATCCATCTCTTCAGCTTCAGCGCGGTATACCAACCCACCCCATGCGCCACGTACT 3900
QY 3901 CAGGCTCATGTACAGTCAAGGAATGTTCTTCTCATCCAACTGCGCCCAATGATG 3960
Db 3901 CAGGCTCATGTACAGTCAAGGAATGTTCTTCTCATCCAACTGCGCCCAATGATG 3960
QY 3961 CTAATGACGACACAGCCACCGCGGTCCAGGCGCGCTCGCTCAAAGTGCACTACAG 4020
Db 3961 CTAATGACGACACAGCCACCGCGGTCCAGGCGCGCTCGCTCAAAGTGCACTACAG 4020
QY 4021 CCCATTCCAGTCTCAGCAACAGCGATTTCCCTTATATGACGCACTTTCAGTCAAGCC 4080
Db 4021 CCCATTCCAGTCTCAGCAACAGCGATTTCCCTTATATGACGCACTTTCAGTCAAGCC 4080
QY 4081 CACCAACCAACAGCAGTGTAAAGGTCCTGAGGAACCGAAAGCCAAATTCCTCTCTC 4140
Db 4081 CACCAACCAACAGCAGTGTAAAGGTCCTGAGGAACCGAAAGCCAAATTCCTCTCTC 4140
QY 4141 CCTTCTACTGCTTCTACCACTGGAGCAGCAAGAACTAGAAATTTTATTTTGT 4200
Db 4141 CCTTCTACTGCTTCTACCACTGGAGCAGCAAGAACTAGAAATTTTATTTTGT 4200
QY 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTCACCTGCGAG 4260
Db 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTCACCTGCGAG 4260

QY 4261 TGAAGATACTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTTCATATTTCCA 4320
Db 4261 TGAAGATACTTGGACCGAGTAGAGGCATTTAGGAACCTTGGGGCTATTTCATATTTCCA 4320
QY 4321 TATCTGTGTTTCAGAGTCCCGCAGGTACCCCAAGCTCTGCTTCCGAAACTGGAAGTTATT 4380
Db 4321 TATCTGTGTTTCAGAGTCCCGCAGGTACCCCAAGCTCTGCTTCCGAAACTGGAAGTTATT 4380
QY 4381 ATTTTAAATAACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAAGT 4440
Db 4381 ATTTTAAATAACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAAGT 4440
QY 4441 GATTCTTCTGCTATTACTGCTAAAAAATAAAAAA 4481
Db 4441 GATTCTTCTGCTATTACTGCTAAAAAATAAAAAA 4481

RESULT 6
ADZ49164
ID ADZ49164 standard; DNA; 4481 BP.
XX AC ADZ49164;
XX AC ADZ49164;
DT 30-JUN-2005 (first entry)
XX Insulin signaling pathway related gene, SEQ ID 493.
DE Antidiabetic; Gene Therapy; Non-insulin dependent diabetes;
KW insulin resistance; gene; ds.
KW Homo sapiens.
OS US2005085436-A1.
PN 21-APR-2005.
PD 08-JUL-2004; 2004US-00887553.
XX 08-JUL-2003; 2003US-0485883P.
PR (LIHH/) LI H.
PA (MAJU/) MA J.
XX Li H, Ma J;
XX WPI; 2005-305194/31.
DR Treating, preventing or ameliorating pathological conditions associated with dysregulation of the insulin signaling pathway (ISP) comprises administering to a subject an amount of a modulator of any of the proteins regulated by ISP.
PS Disclosure; SEQ ID NO 493; 70pp; English.
XX The present invention relates to a method for treating, preventing or ameliorating pathological conditions associated with dysregulation of the insulin signaling pathway (ISP). The method comprises administering to a subject a modulator for ISP-regulated proteins or a pharmaceutical composition comprising the described modulator. The method is useful for treating, preventing or ameliorating pathological conditions associated with dysregulation of the ISP such as type II diabetes or type A syndrome of insulin resistance. The present sequence is a human homolog of a Drosophila gene regulated by ISP. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20050085436.
XX Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

Query Match 100.0%; Score 4481; DB 14; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCGAGAAAGCAACCCAGCGCGCGCCCTCCTCAGCTGTCCCTCCCGGCCCGGG 60
DB 1 ACCCCGAGAAAGCAACCCAGCGCGCGCCCTCCTCAGCTGTCCCTCCCGGCCCGGG 60
QY 61 GCCACTCAGCTTCCTGCTCCGCTCAGCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG 120
DB 61 GCCACTCAGCTTCCTGCTCCGCTCAGCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG 120
QY 121 CACCTCCGCTCCCAACCCGCGCTCCGCGCGCCGCCCTCCGATCGCTCAGCGCCGCA 180
DB 121 CACCTCCGCTCCCAACCCGCGCTCCGCGCGCCGCCCTCCGATCGCTCAGCGCCGCA 180
QY 181 GCTCTCGAGTCCCGGCTGCGCACCCAGTCTCGCGCTTCGCGCAGCCAGGTGGCC 240
DB 181 GCTCTCGAGTCCCGGCTGCGCACCCAGTCTCGCGCTTCGCGCAGCCAGGTGGCC 240
QY 241 GGGTGGCGCTCCCTCAGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
DB 241 GGGTGGCGCTCCCTCAGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 301 CCGGACCGTATCCCTCCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 360
DB 301 CCGGACCGTATCCCTCCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGG 360
QY 361 CAGAGTCCCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGCGCAACGCGCGCGCGG 420
DB 361 CAGAGTCCCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGCGCAACGCGCGCGCGG 420
QY 421 TTTGGCGCGCGCTCCCGCGCGCTCCTTGGTCTCGCGCGCGCTCCCGCGCGCTTCGTCG 480
DB 421 TTTGGCGCGCGCTCCCGCGCGCTCCTTGGTCTCGCGCGCGCTCCCGCGCGCTTCGTCG 480
QY 481 GTCTTCTCCCTCCGCGCGCGCGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 GTCTTCTCCCTCCGCGCGCGCGCGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 CTGCGCGCGCTGCTGCTCCCGCGCGCTTCGCGGTCTCTGCGCGCGCGCGCGCTCCGCG 600
DB 541 CTGCGCGCGCTGCTGCTCCCGCGCGCTTCGCGGTCTCTGCGCGCGCGCGCGCTCCGCG 600
QY 601 TGTTCGCGCGCGCTGCGCGCGCTGATGGCGCGCTCACCATGTCGCTGAAGCCCGAG 660
DB 601 TGTTCGCGCGCGCTGCGCGCGCTGATGGCGCGCTCACCATGTCGCTGAAGCCCGAG 660
QY 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720
DB 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720
QY 721 CAGCGCGCGCGCGCTGCAATGTCGCAAGCCGCGCGCGCGCGCGCTTCTAGCGTCG 780
DB 721 CAGCGCGCGCGCGCTGCAATGTCGCAAGCCGCGCGCGCGCGCGCTTCTAGCGTCG 780
QY 781 CCGCGCGCGCGCTTCGCGCTCCTGCTCGGTCTCTCTGCTTCGCGCGCGCGCGCTCC 840
DB 781 CCGCGCGCGCGCTTCGCGCTCCTGCTCGGTCTCTCTGCTTCGCGCGCGCGCGCTCC 840
QY 841 TCCTCGGTGTCGCGCGACTCCGCGCGCGGAGCGCGCGCTGGCAGAGGTCCGAAC 900
DB 841 TCCTCGGTGTCGCGCGACTCCGCGCGCGGAGCGCGCGCTGGCAGAGGTCCGAAC 900
QY 901 AGTAACAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAAAATATGAG 960
DB 901 AGTAACAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAAAATATGAG 960
QY 961 ATGGTTATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAATGA 1020
DB 961 ATGGTTATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAGTGAATAATGA 1020
QY 1021 GGTATATACAGGAGTTTAAACCTTACAGTCCGAAAGTGTGATTTGGTACTTGTATGCC 1080
DB 1021 GGTATATACAGGAGTTTAAACCTTACAGTCCGAAAGTGTGATTTGGTACTTGTATGCC 1080
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTCCGGGCGCGAAACGTTGAAGAAATATATGGAG 1140

DB 1081 GCACATGAGAAAGTACAGAAATCCAGTTCCGGGCGCGAAACGTTGAAGAAATATATGGAG 1140
QY 1141 ATTTGTTCAAATGTTTACAGCTTTGTTGTGTACAGTTTAAAGATATAGACTCCAGTTAT 1200
DB 1141 ATTTGTTCAAATGTTTACAGCTTTGTTGTGTACAGTTTAAAGATATAGACTCCAGTTAT 1200
QY 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTCAATGGCGAACACAAA 1260
DB 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTCAATGGCGAACACAAA 1260
QY 1261 GAGAAGGACCTGGAGCCCTGGATGCGAGTGAATCAGAGCAATAGAGAACTTGGGCT 1320
DB 1261 GAGAAGGACCTGGAGCCCTGGATGCGAGTGAATCAGAGCAATAGAGAACTTGGGCT 1320
QY 1321 TTGGAATAATGACGTATCTAAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA 1380
DB 1321 TTGGAATAATGACGTATCTAAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA 1380
QY 1381 AATTATGTTGAGTGTCTACGTATGATAGCACTTCTTCGTATACAGTGCCTTAGAA 1440
DB 1381 AATTATGTTGAGTGTCTACGTATGATAGCACTTCTTCGTATACAGTGCCTTAGAA 1440
QY 1441 AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGCAAAACAGTTAGCAAGAA 1500
DB 1441 AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGCAAAACAGTTAGCAAGAA 1500
QY 1501 ATTGAGTCAAGTCCAGTACAAAGCTCAGTGGCCCTGGAAATGATATAGGAGTGA 1560
DB 1501 ATTGAGTCAAGTCCAGTACAAAGCTCAGTGGCCCTGGAAATGATATAGGAGTGA 1560
QY 1561 GAAGAAAATACAGAGAGTTTCAAGAAATTCAGTGAACGTGAGGGGACAGCATATAAC 1620
DB 1561 GAAGAAAATACAGAGAGTTTCAAGAAATTCAGTGAACGTGAGGGGACAGCATATAAC 1620
QY 1621 ACTAGGAAAATAAATATATCTCTCGACAAAGAAATAGAGAGTATATCTCTGGGA 1680
DB 1621 ACTAGGAAAATAAATATATCTCTCGACAAAGAAATAGAGAGTATATCTCTGGGA 1680
QY 1681 AGTGGGAGACAAATTCACCGGTATGGGCCAGCTTGGATCGGGTCCATGCCATCAAGA 1740
DB 1681 AGTGGGAGACAAATTCACCGGTATGGGCCAGCTTGGATCGGGTCCATGCCATCAAGA 1740
QY 1741 TCCACTTCTCAGCTTCAAGTTTCAACCCGAAATCTGTTTCAAGCAAAAGAGTAGTTAAT 1800
DB 1741 TCCACTTCTCAGCTTCAAGTTTCAACCCGAAATCTGTTTCAAGCAAAAGAGTAGTTAAT 1800
QY 1801 GGAGGTGTTCCCTGGCCATCGCTTCCCATCTCTCTCTCTCGCCACCTTCTCGCTAC 1860
DB 1801 GGAGGTGTTCCCTGGCCATCGCTTCCCATCTCTCTCTCTCGCCACCTTCTCGCTAC 1860
QY 1861 CAGTCAGTCCCAACTCTCTTCCACTCGGGCAGCCACCCCTACAGCGCGCCCTCCAGG 1920
DB 1861 CAGTCAGTCCCAACTCTCTTCCACTCGGGCAGCCACCCCTACAGCGCGCCCTCCAGG 1920
QY 1921 CCCCCCTCGCGCCATCCAGACCCCGCTCTACCCCTCTGCTCATGGTCTCCAGTCTCT 1980
DB 1921 CCCCCCTCGCGCCATCCAGACCCCGCTCTACCCCTCTGCTCATGGTCTCTCCAGTCTCT 1980
QY 1981 GTCTCTACTATGCTTAAACGCGATGCTTTCAGAGGGGCTTCCAGGATGTCCCCAAGGCC 2040
DB 1981 GTCTCTACTATGCTTAAACGCGATGCTTTCAGAGGGGCTTCCAGGATGTCCCCAAGGCC 2040
QY 2041 CAGCGACATCTCTGAAATCAGAGTTTCTGCTGGGAGGGTTCATATCCAGTGGCTA 2100
DB 2041 CAGCGACATCTCTGAAATCAGAGTTTCTGCTGGGAGGGTTCATATCCAGTGGCTA 2100
QY 2101 GAAATTTGATCCCAACCCAGTGAAGCAGTACTCTCTCAGTAGCAAGGACAGT 2160
DB 2101 GAAATTTGATCCCAACCCAGTGAAGCAGTACTCTCTCAGTAGCAAGGACAGT 2160
QY 2161 CCTCGGGGGAGCTGCTCATCAGTGTCTAGTGGGTTCCAGATTTATCCCTTAAACT 2220

Db 2161 CCTCGGGGGAACTGTGTCATCATGGTCAGTGGGGTTCAAAGATATCCCTTAAAACT 2220
Qy 2221 CATAGACCCAGGCTCCAGAGACAGAACAGTATGGAAATACCCCCAGTGGGCGAGTCTT 2280
Db 2221 CATAGACCCAGGCTCCAGAGACAGAACAGTATGGAAATACCCCCAGTGGGCGAGTCTT 2280
Qy 2281 GCTTCTCCCAAGCTGGTATATATCCAACTGAAGCTGTGGCATGCCATTCCAGCTGCA 2340
Db 2281 GCTTCTCCCAAGCTGGTATATATCCAACTGAAGCTGTGGCATGCCATTCCAGCTGCA 2340
Qy 2341 TCTCCTAGCGCTGTAGTCCGATCGAACAGAGCTGTACCCCTTCTAGTGGGCTAAA 2400
Db 2341 TCTCCTAGCGCTGTAGTCCGATCGAACAGAGCTGTACCCCTTCTAGTGGGCTAAA 2400
Qy 2401 GATTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGGAATAAAGAAATATTAAA 2460
Db 2401 GATTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGGAATAAAGAAATATTAAA 2460
Qy 2461 CCCAATGAAACATCACCTAGCTTCTCAAAGCTGAAAAAAGATATATCACCGATTGTT 2520
Db 2461 CCCAATGAAACATCACCTAGCTTCTCAAAGCTGAAAAAAGATATATCACCGATTGTT 2520
Qy 2521 TCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAGAAATTTAGGTTACAG 2580
Db 2521 TCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAGAAATTTAGGTTACAG 2580
Qy 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTATAAACAATAAGAGAGGAGAAAA 2640
Db 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTATAAACAATAAGAGAGGAGAAAA 2640
Qy 2641 TCAAGAGATTTGATCAAAAGACAAAATTGAACCAAGTCTAAGGATTTCTTTCATTGAAAAT 2700
Db 2641 TCAAGAGATTTGATCAAAAGACAAAATTGAACCAAGTCTAAGGATTTCTTTCATTGAAAAT 2700
Qy 2701 AGCAGCAGCACTGTACAGTGGGAGCAAGCCGAATAGCCCAAGGATTTCCCTTCA 2760
Db 2701 AGCAGCAGCACTGTACAGTGGGAGCAAGCCGAATAGCCCAAGGATTTCCCTTCA 2760
Qy 2761 ATACTTAGTAAACAGGAGCAAGAGGGGACTGAGGTCACTTCCCAAGGGGTTCCAGACT 2820
Db 2761 ATACTTAGTAAACAGGAGCAAGAGGGGACTGAGGTCACTTCCCAAGGGGTTCCAGACT 2820
Qy 2821 TCCAGCCAGCATGTAAAACAAGAGAAAGACGATAGGAAGAGAAAGACGCACTGAG 2880
Db 2821 TCCAGCCAGCATGTAAAACAAGAGAAAGACGATAGGAAGAGAAAGACGCACTGAG 2880
Qy 2881 CAAGTTAGGAATCAACATTTGAAATCCCAATGCAAGAGGATTCACCCAGCTTCTTCT 2940
Db 2881 CAAGTTAGGAATCAACATTTGAAATCCCAATGCAAGAGGATTCACCCAGCTTCTTCT 2940
Qy 2941 CAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCTCAAGACACACCTAGCCATCT 3000
Db 2941 CAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCTCAAGACACACCTAGCCATCT 3000
Qy 3001 ATGTGGGTGATCAACAGCCAACTCCAGTTTATCTCAGCCTGTGTTGTTTGCACCAAT 3060
Db 3001 ATGTGGGTGATCAACAGCCAACTCCAGTTTATCTCAGCCTGTGTTGTTTGCACCAAT 3060
Qy 3061 ATGATGTATCCAGTCCAGTGCCAGGCTGCAACCTTTATACCCAAATACCTATGACG 3120
Db 3061 ATGATGTATCCAGTCCAGTGCCAGGCTGCAACCTTTATACCCAAATACCTATGACG 3120
Qy 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGCAGTACCAATATGCCCAACAGCGG 3180
Db 3121 CCCATGCCAGTGAATCAAGCCAAAGACATATAGCAGTACCAATATGCCCAACAGCGG 3180
Qy 3181 CAAGACCGAGTCATCAGAGTGCCATGATGACCCAGGTCAGCAGCGGGCCACCGATT 3240
Db 3181 CAAGACCGAGTCATCAGAGTGCCATGATGACCCAGGTCAGCAGCGGGCCACCGATT 3240
Qy 3241 GCAGCCACCCACGAGCTTACTCAGCGAATATGTTGCTTACAGTCTCAGAGTTCCCA 3300
Db 3241 GCAGCCACCCACGAGCTTACTCAGCGAATATGTTGCTTACAGTCTCAGAGTTCCCA 3300

Qy 3301 AATCAGGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
Db 3301 AATCAGGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
Qy 3361 CCTGTAAATACAGGGTAATGCTTAGAATGATGGCAACACAAACACGCGCCAGCTGGTTTA 3420
Db 3361 CCTGTAAATACAGGGTAATGCTTAGAATGATGGCAACACAAACACGCGCCAGCTGGTTTA 3420
Qy 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGATGGATGTATGCAATG 3480
Db 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGATGGATGTATGCAATG 3480
Qy 3481 CCCAAATTTACCATACAAAGGAGACAAAGCCCTTCTTCTACTTTTGCCATTTCCACGGGC 3540
Db 3481 CCCAAATTTACCATACAAAGGAGACAAAGCCCTTCTTCTACTTTTGCCATTTCCACGGGC 3540
Qy 3541 TCCCTTGTCTAGCATGTGGCAACCTTAAACGCTACCTGTGCAACCAATATCTCCACACCT 3600
Db 3541 TCCCTTGTCTAGCATGTGGCAACCTTAAACGCTACCTGTGCAACCAATATCTCCACACCT 3600
Qy 3601 CAGCCTTTCAGTACCCCACTGGACAGCAGCAAAAGCCAAATGTTGGAAGTCACTCTGCA 3660
Db 3601 CAGCCTTTCAGTACCCCACTGGACAGCAGCAAAAGCCAAATGTTGGAAGTCACTCTGCA 3660
Qy 3661 CCCAGTCTCTGTTACAGCACTCAGCACACGAGCCGCCAGGCTCTCCATCTGGCCAGTCCA 3720
Db 3661 CCCAGTCTCTGTTACAGCACTCAGCACACGAGCCGCCAGGCTCTCCATCTGGCCAGTCCA 3720
Qy 3721 CAGCAGCAGTCAAGCAATTTACCGCGGGGTTGGGCAACTCCACCTTCATGACACCT 3780
Db 3721 CAGCAGCAGTCAAGCAATTTACCGCGGGGTTGGGCAACTCCACCTTCATGACACCT 3780
Qy 3781 GCCTCCAAACAGCGAGTGGCCAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTAG 3840
Db 3781 GCCTCCAAACAGCGAGTGGCCAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTAG 3840
Qy 3841 ATCCATCTCTTCACGTTTCAGCGCGGTATACCAACCCACCCACATGGCCACAGTACT 3900
Db 3841 ATCCATCTCTTCACGTTTCAGCGCGGTATACCAACCCACCCACATGGCCACAGTACT 3900
Qy 3901 CAGGCTCATGTACAGTCAAGAAATGGTTCCTTCTCATCAACTGCGCCATGCGCAATGATG 3960
Db 3901 CAGGCTCATGTACAGTCAAGAAATGGTTCCTTCTCATCAACTGCGCCATGCGCAATGATG 3960
Qy 3961 CTAATGAGCAGACAGCCACCGCGGCTCCAGCGCGCTCGCTCAAGTGGCACTACAG 4020
Db 3961 CTAATGAGCAGACAGCCACCGCGGCTCCAGCGCGCTCGCTCAAGTGGCACTACAG 4020
Qy 4021 CCCATTCAGTCTCGACAAACAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAGCC 4080
Db 4021 CCCATTCAGTCTCGACAAACAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAGCC 4080
Qy 4081 CACCAACAAACAGCAGTGTAAAGCTGCGCTGGAGAAACCGAAAGGCGCAAAATTCCTCTCT 4140
Db 4081 CACCAACAAACAGCAGTGTAAAGCTGCGCTGGAGAAACCGAAAGGCGCAAAATTCCTCTCT 4140
Qy 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGCAAAACTAGAAATTTTCAATTTTGTGTTT 4200
Db 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGCAAAACTAGAAATTTTCAATTTTGTGTTT 4200
Qy 4201 TAAATATATATGTTGATTTCTTCTAACTCAATAGAAATGCTAAACAGTTTCACTTGAG 4260
Db 4201 TAAATATATATGTTGATTTCTTGTATCATCTCAATAGGAATGCTAAACAGTTTCACTTGAG 4260
Qy 4261 TGGAGATACTTGGACCGAGTAGAGGCAATTTAGGAACTTGGGGCTATTTCCATAATTTCCA 4320
Db 4261 TGGAGATACTTGGACCGAGTAGAGGCAATTTAGGAACTTGGGGCTATTTCCATAATTTCCA 4320
Qy 4321 TATGCTGTTTTCAGAGTCCCGAGGTACCCAGCTCTGCTTCCCGAAACTGGAAGTTATTT 4380
Db 4321 TATGCTGTTTTCAGAGTCCCGAGGTACCCAGCTCTGCTTCCCGAAACTGGAAGTTATTT 4380

QY 4391 ATTTTAAATACCTTGAAAGTCATGAACATCAGCTAGCAAAAGAAAGTAAACAAGT 4440
 DB 4381 ATTTTAAATACCTTGAAAGTCATGAACATCAGCTAGCAAAAGAAAGTAAACAAGT 4440
 QY 4441 GATTCTGCTGCTATTACTGCTAAAAAAGAAAAA 4481
 DB 4441 GATTCTGCTGCTATTACTGCTAAAAAAGAAAAA 4481

RESULT 7
 AAV30270
 ID AAV30270 standard; DNA; 4367 BP.
 XX
 AC AAV30270;
 XX
 DT 02-OCT-1998 (first entry)
 XX
 DE Gene causative of spinocerebellar ataxia type 2 (SCA2) DNA sequence.
 XX
 KW Spinocerebellar ataxia type 2; SCA2; gene therapy; antisense therapy;
 CAG repeat; neurodegenerative disease; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 49..3990
 FT /tag= a
 FT /product= "Spinocerebellar ataxia type 2 associated
 protein"
 FT repeat_region 544..612
 FT /tag= b
 FT /note= "normal CAG repeat region; this is increased in
 patients with SCA2"
 FT repeat_unit 544..546
 FT /tag= c
 XX
 PN WO9818920-A1.
 XX
 PD 07-MAY-1998.
 XX
 PF 30-OCT-1997; 97WO-JP003946.
 XX
 PR 30-OCT-1996; 96JP-00304059.
 XX
 PA (SRLS-) SRL INC.
 XX
 PI Tsuji S, Sanpei K;
 XX
 DR WPI; 1998-272215/24.
 DR P-PSDB; AAW60213.
 XX
 PT Nucleic acid fragments associated with spinocerebellar ataxia type 2 -
 contain increased number of CAG repeat region compared to normal gene.
 XX
 PS Claim 1; Page 13-22; 38pp; Japanese.
 XX
 CC This represents the sequence of a gene causative of spinocerebellar
 ataxia type 2 (SCA2), a neurodegenerative disease. This gene associated
 with SCA2, has a tri-nucleotide (CAG) repeat region which in the
 CC expression product produces a polyglutamine sequence from Gln-166 to Gln-
 CC 188. In the normal gene there are 15-25 CAG repeats but in SCA2 patients
 CC this number is increased to 35-100. Peptides encoded by nucleic acid
 CC fragments (DNA or RNA) containing sequences from the SCA2 associated
 CC gene, antibodies recognising the peptides and antisense nucleic acids
 CC hybridising with the nucleic acid fragments can be used for the
 CC investigation and diagnosis of SCA2. They can also be used for the
 CC treatment of SCA2 by antisense therapy or gene therapy
 XX
 SQ Sequence 4367 BP; 1124 A; 1328 C; 991 G; 924 T; 0 U; 0 Other;
 Query Match 97.0%; Score 4347.8; DB 2; Length 4367;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 4362; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 115 TATCCGACCTCCCTCCACCCCGCGCTCGGCGCGCGCCCTCCGATCGCTCAGCG 174
 DB 1 TATCCGACCTCCCTCCACCCCGCGCTCGGCGCGCGCCCTCCGATCGCTCAGCG 60
 QY 175 GCGCAGCTCCTCGGAGTCCCGCGGTGGCCACCGAGTCTCCCGCTTCGCGCAGACGAG 234
 DB 61 GCCCAGCTCCTCGGAGTCCCGCGGTGGCCACCGAGTCTCCCGCTTCGCGCAGACGAG 120
 QY 235 TGGCCCGGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGGGGGGGGTGGC 294
 DB 121 TGGCCCGGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGCGGGGGGGGGTGGC 180
 QY 295 GCGGCCCGGAGACGATATCCCTCCCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCC 354
 DB 181 GCGGCCCGGAGACGATATCCCTCCCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCC 240
 QY 355 TCCCGGAGAGCTCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAAGCAAGCGCGCG 414
 DB 241 TCCCGGAGAGCTCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAAGCGCGCGCGCG 300
 QY 415 GCGCGTTTCGCGCGCGCTCCCGCGGTCTTGGTCTCGCGCGCGCTCCCGCGCGCGCTTC 474
 DB 301 GCGCGTTTCGCGCGCGCTCCCGCGGTCTTGGTCTCGCGCGCGCTCCCGCGCGCGCTTC 360
 QY 475 GTGTCGTCTTCTCCCTCCCGCGGTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 534
 DB 361 GTGTCGTCTTCTCCCTCCCGCGGTCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 QY 535 TCCCGCTCGCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 594
 DB 421 TCCCGCTCGCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 480
 QY 595 CCGGCTGTCCCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 654
 DB 481 CCGGCTGTCCCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 540
 QY 655 CCGGCTGTCCCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 711
 DB 541 CCGGCTGTCCCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 600
 QY 712 CAGCAGCAGCAGCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 771
 DB 601 CAGCAGCAGCAGCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 660
 QY 772 CTAGGTCGCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 831
 DB 661 CTAGGTCGCGCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 720
 QY 832 ACGGCTCCCTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 891
 DB 721 ACGGCTCCCTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCTCCCGCGGTCT 780
 QY 892 GGTGAAACAGTAAACAAGGAGTGCCTCAGTCTAGATTTCTTTTGGTGAATCTATGCA 951
 DB 781 GGTGAAACAGTAAACAAGGAGTGCCTCAGTCTAGATTTCTTTTGGTGAATCTATGCA 840
 QY 952 AATATGAGGATGTTTCAATATCTTACATCAGTTTGGTTCCTTCCAAATGTGAAGTACAAGTG 1011
 DB 841 AATATGAGGATGTTTCAATATCTTACATCAGTTTGGTTCCTTCCAAATGTGAAGTACAAGTG 900
 QY 1012 AAAAAATGGAGGTATATGAAGGAGTGTAAAACTTACAGTCCGAAGTGTGATTGGTA 1071
 DB 901 AAAAAATGGAGGTATATGAAGGAGTGTAAAACTTACAGTCCGAAGTGTGATTGGTA 960
 QY 1072 CTTGATGCGGCACATGAGAAAGTACAGATCCAGTTCGGGCGCGGAGAGTGAAGAAATA 1131
 DB 961 CTTGATGCGGCACATGAGAAAGTACAGATCCAGTTCGGGCGCGGAGAGTGAAGAAATA 1020
 QY 1132 ATGGAGAGTATTTTCTTCAAAATGTTTCAAGTCTTGTGGTACAGTTTAAAGATATGGAC 1191
 DB 1021 ATGGAGAGTATTTTCTTCAAAATGTTTCAAGTCTTGTGGTACAGTTTAAAGATATGGAC 1080

QY 1192 TCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTGAATGGC 1251
DB 1081 TCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTGAATGGC 1140
QY 1252 GAACAACAAAGAGAGAGACCTCGAGCCCTGGGATGCGAGTGAACCTCACAGCCAAATGAGGAA 1311
DB 1141 GAAACAANGAGAGAGACCTCGAGCCCTGGGATGCGAGTGAACCTCACAGCCAAATGAGGAA 1200
QY 1312 CTTGAGGCTTTGGAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATAT 1371
DB 1201 CTTGAGGCTTTGGAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATAT 1260
QY 1372 AATGAGNAAATATGTTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTG 1431
DB 1261 AATGAAGAAAATATGTTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTG 1320
QY 1432 CCCTTTAGAAAGAGATAACTCAGAGAATAATTTTAAACGGGAAGCAAGGGCAAAACCAAGTTA 1491
DB 1321 CCCTTAGAAAGAGATAACTCAGAGAATAATTTTAAACGGGAAGCAAGGGCAAAACCAAGTTA 1380
QY 1492 GCAGAAGAAAATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTGGAAAATGATGAT 1551
DB 1381 GCAGAAGAAAATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTGGAAAACGATGAT 1440
QY 1552 AGGAGTGAGGAAGAAAATACACAGCAGTTCAGAGAAATTCAGTGAACGTGAGGGGAC 1611
DB 1441 AGGAGTGAGGAAGAAAATAACACAGCAGTTTACAGAGAAATTCAGTGAACGTGAGGGGAC 1500
QY 1612 AGCATAAACACTAGGGAAAATAAATATATCTCTCTGGACAAAAGAAATAGAGAAAGTCATA 1671
DB 1501 AGCATAAACACTAGGGAAAATAAATATATCTCTCTGGACAAAAGAAATAGAGAAAGTCATA 1560
QY 1672 TCCTGGGAAGTGGGAGACAGAAATTCACGGGTATGGGCCAGCCCTGGATCGGGCTCCATG 1731
DB 1561 TCCTGGGAAGTGGGAGACAGAAATTCACGGGTATGGGCCAGCCCTGGATCGGGCTCCATG 1620
QY 1732 CCATCAAGATCCATTCTCACACTTCAGATTTCAACCCGGAATTCGTGTTAGACAAAGA 1791
DB 1621 CCATCAAGATCCATTCTCACACTTCAGATTTCAACCCGGAATTCGTGTTAGACAAAGA 1680
QY 1792 GTAGTTAATGAGGTGTTCCCTGGGCAATCGCTTGCCCATCTCTCTCTCGGCCACCT 1851
DB 1681 GTAGTTAATGAGGTGTTCCCTGGGCAATCGCTTGCCCATCTCTCTCTCGGCCACCT 1740
QY 1852 TCTCGCTACAGTACAGGTCCCAACTCTCTTCCACTCGGGGAGCCACCCTACAGGCGG 1911
DB 1741 TCTCGCTACAGTACAGGTCCCAACTCTCTTCCACTCGGGGAGCCACCCTACAGGCGG 1800
QY 1912 CCCTCCAGGCCCCCTCGGGGCCATCCAGACCCCGGCTCACCCCTCTGCTCATGTGTTCT 1971
DB 1801 CCCTCCAGGCCCCCTCGGGGCCATCCAGACCCCGGCTCACCCCTCTGCTCATGTGTTCT 1860
QY 1972 CCAGTCTCTGCTCTACTATGCTTAAACGCGATGCTTCAGAAAGGCGCTCCAAGATGTC 2031
DB 1861 CCAGTCTCTGCTCTACTATGCTTAAACGCGATGCTTCAGAAAGGCGCTCCAAGATGTC 1920
QY 2032 CCAAGGGCCACGACATCTCTCGAAATCACAGAGTTTCTGTGGGAGGGGTTCCATATCC 2091
DB 1921 CCAAGGGCCACGACATCTCTCGAAATCACAGAGTTTCTGTGGGAGGGGTTCCATATCC 1980
QY 2092 AGTGGCTTAGAATTTGTATCCCAAAACCCAGTGAAGCAGCTATCTCTCCAGTAGCA 2151
DB 1981 AGTGGCTTAGAATTTGTATCCCAAAACCCAGTGAAGCAGCTATCTCTCCAGTAGCA 2040
QY 2152 AGGACCACTCTCGGGGGGAACGTGGTCACTAGTGGTCACTAGTGGTTCGAAGATTATCC 2211
DB 2041 AGGACCACTCTCGGGGGGAACGTGGTCACTAGTGGTTCGAAGATTATCC 2100
QY 2212 CCTAAAATCTATAGACCCAGGTCTCCACAGACAGAACAGTATGGAAAATACCCCAAGTGGG 2271
DB 2101 CCTAAAATCTATAGACCCAGGTCTCCACAGACAGAACAGTATGGAAAATACCCCAAGTGGG 2160
QY 2272 CCAGTTCCTTGCTTCTCCCAAGCTGGTATTAATTCAACTGAAGCTGTGGCATGCTTAT 2331

DB 2161 CCAGTTCCTTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGCATGCCTATT 2220
QY 2332 CCAGCTGCATCTCTCAGCCCTGCTAGTCTCGCATCGAACAGAGCTGTTTACCCCTTCTAGT 2391
DB 2221 CCAGCTGCATCTCTCAGCCCTGCTAGTCTCGCATCGAACAGAGCTGTTTACCCCTTCTAGT 2280
QY 2392 GAGCTAAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCAGGGAAATAAGAA 2451
DB 2281 GAGCTAAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCAGGGAAATAAGAA 2340
QY 2452 AATATTAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAAGGTATATCA 2511
DB 2341 AATATTAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAAGGTATATCA 2400
QY 2512 CCAGTGTGTTCTGAACTAGAAAACAGATTGATTTTAAAGAAATTAAGAAATGATTTT 2571
DB 2401 CCAGTGTGTTCTGAACTAGAAAACAGATTGATTTTAAAGAAATTTAAGAAATGATTTT 2460
QY 2572 AGTTACAGCCAAAGTTCCTACTCTTCTGAATCTATGGATCAACTACTATAACAAAATAGAGAG 2631
DB 2461 AGTTACAGCCAAAGTTCCTACTCTTCTGAATCTATGGATCAACTACTATAACAAAATAGAGAG 2520
QY 2632 GGAGAAAATCAAGAGATTTGATCAAAAGACAAATTTGAACCAAGTCTAAGGATTTCTTC 2691
DB 2521 GGAGAAAATCAAGAGATTTGATCAAAAGACAAATTTGAACCAAGTCTAAGGATTTCTTC 2580
QY 2692 ATTGAAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAATAGCCCCAGCAT 2751
DB 2581 ATTGAAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAATAGCCCCAGCAT 2640
QY 2752 TCCCTTTCAATACTATTAGTAAACAGCGACAACAAGGGGACCTGAGGTCACTTCCCAAGG 2811
DB 2641 TCCCTTTCAATACTATTAGTAAACAGCGACAACAAGGGGACCTGAGGTCACTTCCCAAGG 2700
QY 2812 GTTCAGACTTCCAGCCAGCATGTAAACAGAGAGAAAGCAGATAGGAAGAGAGAAAGAC 2871
DB 2701 GTTCAGACTTCCAGCCAGCATGTAAACAGAGAGAAAGCAGATAGGAAGAGAGAAAGAC 2760
QY 2872 GCAGCTGAGCAAGTTAGGAAATCAACTTGAATCCCAATGCAAAAGGAGTTCAAACCCAGT 2931
DB 2761 GCAGCTGAGCAAGTTAGGAAATCAACTTGAATCCCAATGCAAAAGGAGTTCAAACCCAGT 2820
QY 2932 TCTTCTCTCAGCCAAAGCCCTTCTACTACGCCAACTTCACTCGGCCCTCAAGCAACCT 2991
DB 2821 TCTTCTCTCAGCCAAAGCCCTTCTACTACGCCAACTTCACTCGGCCCTCAAGCAACCT 2880
QY 2992 AGCCCATCTATGGTGGGTCTATCAACAGCCAACTCCAGTTTATCTCAGCCTGTTGTTTT 3051
DB 2881 AGCCCATCTATGGTGGGTCTATCAACAGCCAACTCCAGTTTATCTCAGCCTGTTGTTTT 2940
QY 3052 GCACCAATATGATGTATCCAGTCCAGTGAGCCAGCGTGCAACCTTTATACCCAATA 3111
DB 2941 GCACCAATATGATGTATCCAGTCCAGTGAGCCAGCGTGCAACCTTTATACCCAATA 3000
QY 3112 CCTATGACGCCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCC 3171
DB 3001 CCTATGACGCCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAAAATATGCC 3060
QY 3172 CAAAGCGGCAAGACAGCATCATCAGATGCCATGATGCACCAGCGTCAGACGGGCG 3231
DB 3061 CAAAGCGGCAAGACAGCATCATCAGATGCCATGATGCACCAGCGTCAGACGGGCG 3120
QY 3232 CCACCGATTGACGACCAACCCACAGCTTACTCCACGCAATATGTTGCTACAGTCCCTCAG 3291
DB 3121 CCACCGATTGACGACCAACCCACAGCTTACTCCACGCAATATGTTGCTACAGTCCCTCAG 3180
QY 3292 CAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCAATTATCAGTCTCAGCATCTCTCAT 3351
DB 3181 CAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCAATTATCAGTCTCAGCATCTCTCAT 3240
QY 3352 GTCTATAGTCTGTAAATACAGGTAATGCTAGAAATGATGGCACCAACACAGCCAG 3411

Db 3241 GTCTATAGTCTGTTAATACAGGGTAATGCTAGATGATGCGACCAACACACAGCCCG 3300
Qy 3412 CTTGGTTTATGATATCTTTTACGAACTCAGTACGGGGCTCATGAGCAGACGATCGATG 3471
Db 3301 CTTGGTTTATGATATCTTTTACGAACTCAGTACGGGGCTCATGAGCAGACGATCGATG 3360
Qy 3472 TATGATGTCCTCAATATACATACAAAGAGAGACAGCCCTTCTTTTACTTTGGCAT 3531
Db 3361 TATGATGTCCTCAATATACATACAAAGAGAGACAGCCCTTCTTTTACTTTGGCAT 3420
Qy 3532 TCCAGGGCTCCTTGTCTCAGCAGATGCGACCTTACGCTACCTGTCACCCACATATCT 3591
Db 3421 TCCAGGGCTCCTTGTCTCAGCAGATGCGACCTTACGCTACCTGTCACCCACATATCT 3480
Qy 3592 CCACACCTTCAGCTTTCAGTACCTCCCTGAGCAGCAGCAAGCAACATGTTGGAAGT 3651
Db 3481 CCACACCTTCAGCTTTCAGTACCTCCCTGAGCAGCAGCAAGCAACATGTTGGAAGT 3540
Qy 3652 CATCTGACCCAGTCTCTGTCAGCACTATCAGCAGCAGCGCCGAGGCTCTCCATCTG 3711
Db 3541 CATCTGACCCAGTCTCTGTCAGCACTATCAGCAGCAGCGCCGAGGCTCTCCATCTG 3600
Qy 3712 GCAGTCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3771
Db 3601 GCAGTCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3660
Qy 3772 ATGACACCTGCTTCCAAACAGCAGTCCGACAGATAGTTCCTGAGCAGCAGCAACAGACT 3831
Db 3661 ATGACACCTGCTTCCAAACAGCAGTCCGACAGATAGTTCCTGAGCAGCAGCAACAGACT 3720
Qy 3832 GTCTTTACGATCATCTTCTACGTTTACGTTTACGTTTACGTTTACGTTTACGTTTACGTT 3891
Db 3721 GTCTTTACGATCATCTTCTACGTTTACGTTTACGTTTACGTTTACGTTTACGTTTACGTT 3780
Qy 3892 CAGTACCTCAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3951
Db 3781 CAGTACCTCAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3840
Qy 3952 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4011
Db 3841 CCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3900
Qy 4012 GCACTACAGCCCATTCAGTCTCGACACAGCAGCAGTTCCTTATATGACGACACCTTCA 4071
Db 3901 GCACTACAGCCCATTCAGTCTCGACACAGCAGCAGTTCCTTATATGACGACACCTTCA 3960
Qy 4072 GTACAGCCACACCAACAGCAGTGTGAGGCTGCTGAGGAGCAGGAGCAGGAGCAGGAGC 4131
Db 3961 GTACAGCCACACCAACAGCAGTGTGAGGCTGCTGAGGAGCAGGAGCAGGAGCAGGAGC 4020
Qy 4132 TCCCTCTCCTCTCTACTGCTTCTACCACTGGAAGCAGACAGAACTAGAAATTTCAATTTA 4191
Db 4021 TCCCTCTCCTCTCTACTGCTTCTACCACTGGAAGCAGACAGAACTAGAAATTTCAATTTA 4080
Qy 4192 TTTTGTGTTTAAATATATATGTTGATTTCTTGTAAATCCAAATGGAATGCTTAAAGTT 4251
Db 4081 TTTTGTGTTTAAATATATATGTTGATTTCTTGTAAATCCAAATGGAATGCTTAAAGTT 4140
Qy 4252 CACTTGCAGTGGAGATGATCTTGCACCGAGTATGAGGATTTAGGAATTTGGGGCTATTCC 4311
Db 4141 CACTTGCAGTGGAGATGATCTTGCACCGAGTATGAGGATTTAGGAATTTGGGGCTATTCC 4200
Qy 4312 ATAAATTCATATGCTGTTTCCAGAGTCCCGAGGTACCCAGCTCTGCTTCCGAAACTGG 4371
Db 4201 ATAAATTCATATGCTGTTTCCAGAGTCCCGAGGTACCCAGCTCTGCTTCCGAAACTGG 4260
Qy 4372 AAGTATTTATTTTAAATTAACCTTGAAGTATGAAACATCATGAAACATCATGAAAGT 4431
Db 4261 AAGTATTTATTTTAAATTAACCTTGAAGTATGAAACATCATGAAACATCATGAAAGT 4320
Qy 4432 AACAGAGTGAATCTTCTGCTCTATCTGCTTAAATTAACCTTGAAGTATGAAAGT 4478
Db 4321 AACAGAGTGAATCTTCTGCTCTATCTGCTTAAATTAACCTTGAAGTATGAAAGT 4367

RESULT 8
AAT78912
ID AAT78912 standard; cDNA; 4200 BP.
XX
AC AAT78912;
XX
DT 09-FEB-1998 (first entry)
XX
DE Spinocerebellar ataxia gene SCA2.
KW Monoclonal antibody; neurodegenerative disease; polyglutamine; TBP;
KW repeat region; affinity; TARA binding protein; Kennedy disease;
KW transcription initiation factor; lymphoblastic cell line; schizophrenia;
KW Huntington's disease; dominant autosomal spinocerebellar ataxia;
KW X-linked spino-bulbar muscular atrophy; familial spastic paraplegia;
KW dentatorubral-pallidolusial atrophy; bipolar affective disorder;
KW manic depressive psychosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_feature 1..3986
FT /tag= g
FT /note= "sequence contained in DAN1 clone"
FT CDS 3..2747
FT /tag= a
FT /product= "SCA2 protein"
FT /note= "this CDS contains a putative translational start
FT codon for the SCA2 protein at positions 243-245"
FT CDS 3..242
FT /tag= c
FT /note= "putative open reading frame which is in frame
FT with the putative translational start site of the SCA2
FT open reading frame"
FT misc_signal 239..245
FT /tag= d
FT repeat_region /note= "putative Kozak consensus signal"
FT 258..323
FT /tag= e
FT /note= "encodes polyglutamine repeat region; contains
FT repeats of CAG with 2 CAA codons interspersed"
FT repeat_unit 258..260
FT /tag= f
FT /note= "CAG repeats"
FT CDS 2594..3640
FT /tag= b
FT /note= "this second open reading frame may be derived by
FT a frameshift or by alternative splicing"
FT 3987..4200
FT /tag= h
FT /note= "derived from the EST's AAH92640, AAN90240 and
FT AAZ13574 from dbEST database"
FT 4023..4029
FT /tag= i
FT /note= "region which differs in length between the
FT sequences of the EST clones AAH92640, AAN90240 and
FT AAZ13574"
XX
PN W09717445-A1.
XX
PD 15-MAY-1997.
XX
PF 08-NOV-1996; 96WO-FR001773.
XX
PR 10-NOV-1995; 95FR-00013576.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Tora L, Lutz Y, Trottier Y, Mandel J;
XX

DR WPI: 1997-281034/25.
DR P-PSDB; AAW24800, AAW24801.
XX
PT Antibody 1C2 used for treating or preventing neuro-degenerative diseases
PT - associated with proteins containing long poly:glutamine repeats, e.g.
PT Huntington's disease.
XX
PS Claim 21; Page 45-47; 69pp; French.
XX
CC The invention relates to a monoclonal antibody (Mab) 1C2 for the
CC treatment of neurodegenerative diseases associated with the presence of
CC polyglutamine repeat regions. This Mab is already known for its affinity
CC to the TAR4 binding protein (TBP) transcription initiation factor,
CC especially at the amino acid sequence LEEQRQOQQQ found at the N-
CC terminus of TBP. Mab 1C2 has been shown to have a high affinity for
CC polyglutamine repeats with a proportional affinity to the number of
CC glutamine repeats. This affinity has been used to identify genes encoding
CC proteins containing long polyglutamine repeats which are implicated in
CC neurodegenerative diseases. A screen of an expression library, generated
CC from a lymphoblastic cell line from a patient suffering from
CC spinocerebellar ataxia (SCA), with Mab 1C2 isolated 6 new sequences
CC (AAAT7896-T78911) encoding polyglutamine repeats. Mab 1C2 also isolated
CC the complete SCA2 gene in clone DAN1 (sequence presented here). The
CC sequence appears to contain 2 open reading frames (ORF) the second of
CC which may be generated by an frameshift slippage or by an alternative
CC splicing event. The first ORF also encodes a 22 amino acid polyglutamine
CC repeat region near the N-terminus of the protein. Normal SCA2 alleles
CC contain 17-29 CAG triplet repeats with 1-3 CAA repeats interspersed
CC whereas the mutant sequence from patients with SCA contains at least 30,
CC preferably 37-50 CAG repeats. Mab 1C2, active fragment of it or nucleic
CC acids encoding it are specifically used to treat Huntington's disease,
CC SCA types 1-5 or 7, X-linked spino-bulbar muscular atrophy (Kennedy
CC disease), dentatorubral-pallidolusial atrophy, dominant autosomal
CC spinocerebellar ataxia, familial spastic paraplegia, bipolar affective
CC disorder, manic depressive psychoses and schizophrenia
XX
SQ Sequence 4200 BP; 1152 A; 1200 C; 913 G; 935 T; 0 U; 0 Other;
Query Match 88.1%; Score 3949; DB 2; Length 4200;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 4020; Conservative 0; Mismatches 0; Indels 61; Gaps 1;
QY 401 ACGGCAACGGCGGCGGCGGTTTCGGCCCGGCTCCGGCGGCTCTCGGTTCGGCGGCG 460
DB 1 ACGGCAACGGCGGCGGCGGTTTCGGCCCGGCTCCGGCGGCTCTCGGTTCGGCGGCG 60
QY 461 CTCGCCGCGCCCTTCGTGTCGTCCTTCCTCCCGTCGCGAGCCCGGCGCCCTCCGGCGG 520
DB 61 CTCGCCGCGCCCTTCGTGTCGTCCTTCCTCCCGTCGCGAGCCCGGCGCCCTCCGGCGG 120
QY 521 CGCCAACCCGCGCCTCCCGGCTCGCGCGCCGTCGTCCTCCCGCGGTTCCGGCGTCTCCT 580
DB 121 CGCCAACCCGCGCCTCCCGGCTCGCGCGCCGTCGTCCTCCCGCGGTTCCGGCGTCTCCT 180
QY 581 TGGCGCGCGCGGCTCCCGGCTGTCCTCCCGCGCGGTCGAGCGGTCGTCGTCGTCGTCGTC 640
DB 181 TGGCGCGCGCGGCTCCCGGCTGTCCTCCCGCGCGGTCGAGCGGTCGTCGTCGTCGTCGTC 240
QY 641 CCATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 700
DB 241 CCATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300
QY 701 AGC 760
DB 301 AGC 360
QY 761 GCAGCGGCTTCTAGCTGTCGCGCGCGCGCGCTTCGCGCGCTCTGTCGTCGTCGTCGTC 820
DB 361 GCAGCGGCTTCTAGCTGTCGCGCGCGCGCGCTTCGCGCGCTCTGTCGTCGTCGTCGTC 420
QY 821 GGTCTTCGGCCAGCGGCTCCCTCTCGTGTGTCGCGCGGAGCCTTCGCGCGGCGGAGGCGCG 880
DB 421 GGTCTTCGGCCAGCGGCTCCCTCTCGTGTGTCGCGCGGAGCCTTCGCGCGGCGGAGGCGCG 480

QY 1961 CTCTGGTCTCCAGCTCCTGTCTCTACTATGTCCTAAACGATGCTTTCAGAGGGCTC 2020
DB |||||
QY 1561 CTCTGGTCTCCAGCTCCTGTCTCTACTATGTCCTAAACGATGCTTTCAGAGGGCTC 1620
DB |||||
QY 2021 CAAGGATGTCCTCCAAAGGCCAGCGACATCTCTCGAAATCAAGAGTTCCTGCTGGAGGG 2080
DB |||||
QY 1621 CAAGGATGTCCTCCAAAGGCCAGCGACATCTCTCGAAATCAAGAGTTCCTGCTGGAGGG 1680
DB |||||
QY 2081 GTTCCATATCCAGTGGCTAGATTTGTATTCCTCCAAACCCAGCCAGTGAAGCAGTACTC 2140
DB |||||
QY 1681 GTTCCATATCCAGTGGCTAGATTTGTATTCCTCCAAACCCAGCCAGTGAAGCAGTACTC 1740
DB |||||
QY 2141 CTCCAGTACAGGACAGTCTCCTCGGGGGAACGTGTCTACGTGTCTAGTGGGTTTC 2200
DB |||||
QY 1741 CTCCAGTACAGGACAGTCTCCTCGGGGGAACGTGTCTACGTGTCTAGTGGGTTTC 1800
DB |||||
QY 2201 CAAGATTTATCCCTAAAACTCATAGACCCAGGTCTCCAGAGACAGAACTAGTTGGAATA 2260
DB |||||
QY 1801 CAAGATTTATCCCTAAAACTCATAGACCCAGGTCTCCAGAGACAGAACTAGTTGGAATA 1860
DB |||||
QY 2261 CCCCAGTGGGCGAGTCTTGTCTTCTCCCAAGCTGGTATTTATTCAACTGAAGCTGTG 2320
DB |||||
QY 1861 CCCCAGTGGGCGAGTCTTGTCTTCTCCCAAGCTGGTATTTATTCAACTGAAGCTGTG 1920
DB |||||
QY 2321 CCATGCTATTCAGTGTGATCTCTAGCTGTCTAGTCTCTGCTGATCGAACAGAGTGTTA 2380
DB |||||
QY 1921 CCATGCTATTCAGTGTGATCTCTAGCTGTCTAGTCTCTGCTGATCGAACAGAGTGTTA 1980
DB |||||
QY 2381 CCCCCTTCTAGTCAGGCTAAAGATTCAGGCTTCAAGATCAGAGGAGAACTCTCTGCGAG 2440
DB |||||
QY 1981 CCCCCTTCTAGTCAGGCTAAAGATTCAGGCTTCAAGATCAGAGGAGAACTCTCTGCGAG 2040
DB |||||
QY 2441 GGAATAAGAAATATTTAAACCAATGAACCAATCATCTAGTCTTCTCAAAAGCTGAAACA 2500
DB |||||
QY 2041 GGAATAAGAAATATTTAAACCAATGAACCAATCATCTAGTCTTCTCAAAAGCTGAAACA 2100
DB |||||
QY 2501 AAGGTATATCAAGTGTGTTCTGAACATAGAAACAGATTTGATTTTAAAGAAATTTA 2560
DB |||||
QY 2101 AAGGTATATCAAGTGTGTTCTGAACATAGAAACAGATTTGATTTTAAAGAAATTTA 2160
DB |||||
QY 2561 AGAATGATTTTAGTTTACAGCAAGTCTTACTTCTGAATCTATGATCAATCTATAACA 2620
DB |||||
QY 2161 AGAATGATTTTAGTTTACAGCAAGTCTTACTTCTGAATCTATGATCAATCTATAACA 2220
DB |||||
QY 2621 AAAATAGAGAGGGAGAAATCAAGAGATTTGATCAAAAGCAAAATGAAACCAAGTGCTA 2680
DB |||||
QY 2221 AAAATAGAGAGGGAGAAATCAAGAGATTTGATCAAAAGCAAAATGAAACCAAGTGCTA 2280
DB |||||
QY 2681 AGAATTTCTTTCATTTGAAATAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATA 2740
DB |||||
QY 2281 AGAATTTCTTTCATTTGAAATAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATA 2340
DB |||||
QY 2741 GCCCCAGCATTTCCCTTCAATCTTACTTGTAAACAGGAGCAGAGGGGACCTGAGGTCA 2800
DB |||||
QY 2341 GCCCCAGCATTTCCCTTCAATCTTACTTGTAAACAGGAGCAGAGGGGACCTGAGGTCA 2400
DB |||||
QY 2801 CTTTCCAAAGGGTTTACAGCTTCAGGCCAGCATGTAAACAGAGAAACGATTAAGGAAG 2860
DB |||||
QY 2401 CTTTCCAAAGGGTTTACAGCTTCAGGCCAGCATGTAAACAGAGAAACGATTAAGGAAG 2460
DB |||||
QY 2861 AGAAGAAAGACGAGCTGAGCAAGTTTGAATCAACATTTGAATCCCAATGCAAGGAGT 2920
DB |||||
QY 2461 AGAAGAAAGACGAGCTGAGCAAGTTTGAATCAACATTTGAATCCCAATGCAAGGAGT 2520
DB |||||
QY 2921 TCAACCCAGTTCCTTCTCTCAGCAGAAAGCTTCTACTACCCCACTTCACTCGGCTC 2980
DB |||||
QY 2521 TCAACCCAGTTCCTTCTCTCAGCAGAAAGCTTCTACTACCCCACTTCACTCGGCTC 2580
DB |||||
QY 2981 AAGCAACACTAGCCCATCTATGCTGGGTCTACACAGGCAACTCCAGTTTATCTCAGC 3040
DB |||||
QY 2581 AAGCAACACTAGCCCATCTATGCTGGGTCTACACAGGCAACTCCAGTTTATCTCAGC 2640
DB |||||
QY 3041 CTGTTTTTTTGCCAAAATATGATGTATTCAGTCCCGAGTGAGCCCGAGGGCTGCAACCTTT 3100
DB |||||

DB 2641 CTGTTTTTTTGCCAAAATATGATGTATCCAGTCCCGAGCCAGCGCTGCAA----- 2696
QY 3101 TATACCAATACCTATGACGCCCATGCCAGTGAATCAAGCCAAAGACATATAGACAGTAC 3160
DB -----TAC 2699
QY 3161 CAAATATGCCCAACAGCGGCAAGACAGCATCATCAGATGCCATGATGACCCAGCGT 3220
DB CAAATATGCCCAACAGCGGCAAGACAGCATCATCAGATGCCATGATGACCCAGCGT 2759
QY 3221 CAGCAGCGGGCCCAACCGATTGCAGCCACCCACAGCTTACTTCCACGAAATATGTGCT 3280
DB CAGCAGCGGGCCCAACCGATTGCAGCCACCCACAGCTTACTTCCACGAAATATGTGCT 2819
QY 3281 ACAGTCTCAGCAGTTCCAAAATCAGGCCCTTGTTCAGCATGTGCCATATCAGTCTC 3340
DB ACAGTCTCAGCAGTTCCAAAATCAGGCCCTTGTTCAGCATGTGCCATATCAGTCTC 2879
QY 3341 AGCATCTCATGTCTATAGTCTCTGTATACAGGGTAACTAGATGATGGCACCACAA 3400
DB AGCATCTCATGTCTATAGTCTCTGTATACAGGGTAACTAGATGATGGCACCACAA 2939
QY 3401 CACAGCCCGAGCTGTTTGTAGTATCTTCTCAGCAACTCAGTACGGGGCTCATGAGCAGA 3460
DB CACAGCCCGAGCTGTTTGTAGTATCTTCTCAGCAACTCAGTACGGGGCTCATGAGCAGA 2999
QY 3461 CGCATGCGATGTATGATGTCCAAATTCATACAAAGAGAGACAAGCCCTTCTTCT 3520
DB CGCATGCGATGTATGATGTCCAAATTCATACAAAGAGAGACAAGCCCTTCTTCT 3059
QY 3521 ACTTTGCCATTTCCAGGGGCTCCCTTGTCTAGCAGTATGCCACCTAACGCTACCTGCG 3580
DB ACTTTGCCATTTCCAGGGGCTCCCTTGTCTAGCAGTATGCCACCTAACGCTACCTGCG 3119
QY 3581 ACCCATACTCTCCACACCTCAGCCTTACGCTTACCCCACTGGACAGCAAGAAAGCCAA 3640
DB ACCCATACTCTCCACACCTCAGCCTTACGCTTACCCCACTGGACAGCAAGAAAGCCAA 3179
QY 3641 ATGTTGGAAGTCACTCTGACCCAGTCTGTTTTCAGCAGCATCAGCAGCGCGCCAGG 3700
DB ATGTTGGAAGTCACTCTGACCCAGTCTGTTTTCAGCAGCATCAGCAGCGCGCCAGG 3239
QY 3701 CTCTCATCTGGCAGTCCAGCAGCAGTACGCTTACCCAGCGGGCTTGGGCCAA 3760
DB CTCTCATCTGGCAGTCCAGCAGCAGTACGCTTACCCAGCGGGCTTGGGCCAA 3299
QY 3761 CTTCCACCTTCCATGACCTGCTCCACACAGCAGTCCGCAAGATAGTTTCCAGCAG 3820
DB CTTCCACCTTCCATGACCTGCTCCACACAGCAGTCCGCAAGATAGTTTCCAGCAG 3359
QY 3821 CACACAGACTGTCTTTAGTCCATCTTCTCAGTTTTCAGCGGCTATACCAACCCAC 3419
DB CACACAGACTGTCTTTAGTCCATCTTCTCAGTTTTCAGCGGCTATACCAACCCAC 3940
QY 3881 CCCCAGTCCCAAGTCACTCAGGCTCATGTATCAGTACGGAATGTTTCTTCTCATCCAA 3479
DB CCCCAGTCCCAAGTCACTCAGGCTCATGTATCAGTACGGAATGTTTCTTCTCATCCAA 4000
QY 3941 CTTGCCCATGCGCAATGATGTAAATGACACAGCAGCCCGGCTCCCGAGCGGCC 3539
DB CTTGCCCATGCGCAATGATGTAAATGACACAGCAGCCCGGCTCCCGAGCGGCC 4060
QY 4001 TCGCTCAAAAGTGCATCTACAGCCCATTCAGTCTTCGACAAACAGCGCATTTCCCTATATGA 3540
DB TCGCTCAAAAGTGCATCTACAGCCCATTCAGTCTTCGACAAACAGCGCATTTCCCTATATGA 4061
QY 4061 CGCACCTTTTCAAGAGCCCAACCAAGAGTGTAAAGCTGCCCTGGAGGAACCG 3659
DB CGCACCTTTTCAAGAGCCCAACCAAGAGTGTAAAGCTGCCCTGGAGGAACCG 4120
QY 4121 AAAGGCCAAATTCCTCTCTCTTCTACTGCTTCTACCAACTGGAAGACAGAAACTAG 4180
DB |||||

Db 3660 AAAGCCAAATCCCTCCTCCTCCTCTACTGCTTCTACCAACTGGAAGCACAGAAACCTAG 3719
Qy 4181 AATTTCATTATTTTGTGTTTTTAAATAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAA 4240
Db 3720 AATTTCATTATTTTGTGTTTTTAAATAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAA 3779
Qy 4241 TGCTAACAGTTCACCTGAGTGGAAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACTTG 4300
Db 3780 TGCTAACAGTTCACCTGAGTGGAAAGATACCTTGGACCGAGTAGAGGCATTTAGGAACTTG 3839
Qy 4301 GGGGCTATTCCATAATTCATATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTT 4360
Db 3840 GGGGCTATTCCATAATTCATATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTT 3899
Qy 4361 GCCGAACTGGAAGTATTTATTTTAAATACCTTTGAAAGTTCATGAACACATCAGCTA 4420
Db 3900 GCCGAACTGGAAGTATTTATTTTAAATACCTTTGAAAGTTCATGAACACATCAGCTA 3959
Qy 4421 GCAAAAGAGTAACCAAGAGTGATTTCTTGCTGCTATTACTGCTATAAAAAA 4480
Db 3960 GCAAAAGAGTAACCAAGAGTGATTTCTTGCTGCTATTACTGCTATAAAAAA 4019
Qy 4481 A 4481
Db 4020 A 4020

RESULT 9
ID ABV75290 standard; DNA; 4225 BP.
XX AC ABV75290;
XX DT 07-MAR-2003 (first entry)
XX DE Mouse SCA2 polypeptide encoding DNA.
XX KW SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity;
KW memory impairment; mutant; mouse; gene; ds.
XX OS Mus musculus.
XX FH Key Location/Qualifiers
FT CDS 27..3884
FT FT /*tag= a
FT FT /transl_except= (pos: 222..224, aa: Xaa)
FT FT /note= "Xaa is Pro"
FT FT /product= "SCA2"
FN WO200289567-A2.
XX PD 14-NOV-2002.
XX PF 07-MAY-2002; 2002WO-US014698.
XX PR 07-MAY-2001; 2001US-0289231P.
XX PA (CEDA-) CEDARS SINAI MEDICAL CENT.
XX PI Pulst SM;
XX DR WPI; 2003-103467/09.
XX DR P-PSDB; ABB82697.
XX PT New mutant mouse, useful as a disease model for identifying agents for
PT treating obesity or memory impairment, comprises a disrupted
PT spinocerebellar ataxia 2 gene.
XX FS Disclosure; Page 56-62; 76pp; English.
XX CC The invention relates to a mutant non-human mammal having a disrupted
CC spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a
CC disease model for identifying agents or compounds that may treat obesity

CC or memory impairment. The stem cell is used for generating the mutant
CC animal. The present sequence represents the nucleotide sequence (GenBank
XX Accession No. AF041472) encoding the mouse SCA2 polypeptide
SQ Sequence 4225 BP; 1007 A; 1324 C; 1042 G; 851 T; 0 U; 1 Other;
Query Match 70.0%; Score 3138.2; DB 10; Length 4225;
Best Local Similarity 86.2%; Pred. No. 0;
Matches 3662; Conservative 1; Mismatches 473; Indels 112; Gaps 14;
Qy 222 CGCGCAGCAGGTGGCCGGTGGCGCTCGCTCAGGGCGGGCGGGCGGCGGCG 281
Db 71 CCCGAGCGCGCCCGCGGGCTGGGCGCGCGGCTCGCTCCGCGAGCGGCGCG 130
Qy 282 GGGCGGCGGTGGCGCGGCCCGCGGACCGTATCCCTCCGCCGCCCTCCCGCGCGGCC 341
Db 131 GCGCGGGGGGGGGGGGGGGGGTGGGTATCCCTCCGCCGCCCTCCCGCGGGGCC 190
Qy 342 CGGCCCGCCCTCCCTCCGGCAGAGCTCGCTCCCTCCGCCCTCAGACTGTTTGTAGCAA 401
Db 191 CGGCGCCCTCCCGCGGGCGCGCTCGCCACCSTGCGCTCAGACTGTTTGTAGCAA 250
Qy 402 CGGCAACGGCGGGCGGGGTTTCGGCGCGGCTCCGGCGGGCTCCTTGTCTCGGCGGGCC 461
Db 251 CGGCCA-----CGGCGGCTCCCGCGGGCTCCCGCGGGCTGCTCGGTGTCTCGGGGCC 304
Qy 462 TCCCGCGCCCTTCTGTCGTCCTTCTCCCTCCCGAGCCCGGGCGGCCCTCCGGCGCG 521
Db 305 TCCCGCGCCCTTCTGTCGTTGCTGTCGCTCTG-----GCCCGCGGGCCAC 352
Qy 522 GCCAACCGCGCCCTCCCGCTCGCGCGCGTGGTCCCGCGCGCTTCCGGCGTCTCCTT 581
Db 353 GCGCGCGCGCGCTCGCGCGCGGTCGCGCGCTCGCGCGCTCGCGCGCTCGCTCTC 412
Qy 582 ---GGGGCGCGCGCTCCGGCTGTCCCGCGCGGGTGGCGCGGTGTATGGGCCCTT 638
Db 413 CTCGCGCGCGCGCGCTGTCGCGCGCGCGCTTCCCGCGCGGGTGTATGGGCCCTT 472
Qy 639 CACCATGTGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 698
Db 473 CACCATGTGCTGAAGC----- 490
Qy 699 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 758
Db 491 -----GCAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 532
Qy 759 CGGCGCGCGCTTCTAGGTCGCGCGCGCGCGCTTCCCGCGCGGGTGTATGGGCCCTT 818
Db 533 CGGCGGGCTGCTCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGGGCTTC 592
Qy 819 CTCGTCCTCGGCGCACGGCTCCTCTCGGTG---GTGCGCGCGCACCTCCGGCGCGGGAG 875
Db 593 CGTGTGCGCGCGCGCGCGCGCGCGGTGGGTCTTCTCGCGCGCGCGCGCGCGCGCGCG 652
Qy 876 GCCCGGCTGGGCGAGGTGAAACAGTAACAAAGGACTGCTCAGTCTACGATTTCTTT 935
Db 653 TCCCGGCTGGGCGAGGTGGAACAGTAGCAAGGACTGCTCAGCTACGATTTCTTT 712
Qy 936 TGATGGAATCTATGCAAAATATGAGATGGTTTCATATATTACATCAGTTGTTGGCTCAA 995
Db 713 TGATGGAATCTATGCAAAACGTGAGGATGGTTTCATATATTACGTAGTTGTTGGATCGAA 772
Qy 996 ATGTGAAGTACAAAGTGAAGATATATATGAAGAGTGTATTAAGAGTGTATTAAGAGTCC 1055
Db 773 ATGTGAAGTACAAAGTGAAGATATATGAAGAGTGTATTAAGAGTGTATTAAGAGTCC 832
Qy 1056 GAAGTGTGATTTGGTACTTGTATGCCGACATGAGAAAAGTACAGAAATCCAGTTCCGGGGCC 1115
Db 833 TAAGTGTGACTTGGTACTTGTATGTCGACATGAGAAAAGTACAGAAATCCAGTTCCGGGGCC 892
Qy 1116 GAAACGTGAAGAAATATGAGAGTATTTGTTTCAATGTTACAGACTTTGTTGGGTACA 1175
Db 893 AAAACGTGAAGAAATATGAGAGTGTGTTTGTTCAAATGCTCAGACTTCGTTGTTGGTACA 952

QY 885 GGGCAGAGGTCAAACAGTAAACAAAGGACTGCTCCTCAGTCTACGATTTCTTTTGATGGAAT 944
DB |||||
QY 235 GGGCAGAGGTGGAAACAGTAGCAAGGAGCTGCTCAGCCTACGATTTCTTTTGATGGAAT 294
DB |||||
QY 945 CTATCAAAATATGAGGATGTTTCATATACCTACATCAGTTGTTGGCTCCAAATGGAAGT 1004
DB |||||
QY 295 CTATGCAACGTGAGATGTTTCATATACCTACGTCAGTTGTTGGATCGAAATGGAAGT 354
DB |||||
QY 1005 ACAAGTGAATAATGGAGGTATATGAAGGAGTTTTTAAATCTACAGTCCGAAGTGTGA 1064
DB |||||
QY 355 ACAAGTGAATAACGAGGCATATATGAAGGAGTTTTTAAACATACAGTCTTAAGTGTGA 414
DB |||||
QY 1065 TTTGGTACTTTGATGCGGCACATGAGAAAGTACAGATCCAGTTGCGGGCCGAAACGTGA 1124
DB |||||
QY 415 CTTGGTACTTTGATGCTGCACATGAGAAAGTACAGAACTCCAGTTGCGGGCCAAACGTGA 474
DB |||||
QY 1125 AGAAATATGGAGAGTATTTTCTCAAAATGTTTCAGACTTTGTTGGTACAGTTTAAAGA 1184
DB |||||
QY 475 AGAAATATGGAGAGTGTGTTTCTCAAAATGCTCAGACTTCGTTGTTGGTACAGTTTAAAGA 534
DB |||||
QY 1185 TATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGT 1244
DB |||||
QY 535 TACAGACTCCAGTTATGCAAGGAGATGCTTTTACTGACTCTGCTCTCAGCGCAAGT 594
DB |||||
QY 1245 GAATGCGAACACAAAGAGAAGGACTGGAGCCCTGGGATGAGGTGAACCTCACAGCCAA 1304
DB |||||
QY 595 GAATGCTGAGCACAAAGGAGAAGGACTGGAGCCCTGGGATGAGGAGGACTCACGGCCAG 654
DB |||||
QY 1305 TGAAGNACTTTGAGGCTTTTGGAAATGACGTATCTAATGATGGATCCCAATGATATGTT 1364
DB |||||
QY 655 CGAGGAGCTGGAG---CTGGAGAATGATGTCTAATGGATGGGACCCCAATGACATGTT 711
DB |||||
QY 1365 TCGATATAATGAAGAAATATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCGTA 1424
DB |||||
QY 712 TCGATATAATGAAGAAATATGTTGTTAGTGTGTCTCCATATGATAGCAGTTTATCTTCAT 771
DB |||||
QY 1425 TACAGTGCCTTTAGAAAGAGATAAATCAGAGAAATTTTTTAAACGGGAAGCAGGCAAA 1484
DB |||||
QY 772 TACGGTTCTTTAGAAAGGACAACTCAGAGAAATTTCTTAAACGGGAGGCAAGGCAAA 831
DB |||||
QY 1485 CCAGTTAGCAGAGAAATTTGAGTCAAGTGCACAGTACAAAGCTCCGAGTGGCCCTGGAAA 1544
DB |||||
QY 832 CCAGTTAGCAGAGAAATTTGAATCCAGTGTCTCAGTACAAAGCTCGTGTGCGCCCTTGAGAA 891
DB |||||
QY 1545 TGAATGACAGTACAGGAGAAATAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGA 1604
DB |||||
QY 892 TGAATGACGAGTACAGGAGAAATAATACACAGCAGTCCAGAGNAATCTGAGTGACCGGA 951
DB |||||
QY 1605 GGGGCA CAGCATAAACA CACTAGGGAATAAATATATTCCTCTGGACAAAGAAATAGAGA 1664
DB |||||
QY 952 GGGGCATGCCCCAACACTAGGGA CAATAAATATATTCCTCTGGACAAAGAAACAGAGA 1011
DB |||||
QY 1665 AGTCATATCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGGCGCAGCTGGATCGGG 1724
DB |||||
QY 1012 AGTCCTATCTCTGGGGAAGTGGGAGACAGAGCTCACACGAGATGGGCGAGCTGGGCCAGG 1071
DB |||||
QY 1725 CTCATGCCATCAAGATCCACTTCTCACA CTTCAGATTTTCAACCCGAAATTCGTTTCAGA 1784
DB |||||
QY 1072 CTCATGCCGTCAAGAGCTGCTTCTCACACTTCAGATTTTCAACCCGAAACGCTGGCTCAGA 1131
DB |||||
QY 1785 CCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCTTGGCCATCTCTCTCTCTCG 1844
DB |||||
QY 1132 CCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCTTGGCCATCTCTCTCTCTCG 1191
DB |||||
QY 1845 CCACACTTCTCGTACAGTTCAGGTCCAACTCTCTTCACACTCGGGGACGCAACCCCTAC 1904
DB |||||
QY 1192 CCCACCTTCTCGTACAGTTCAGGTCCCACTCTCTTCCACCTCGGGCAGCCACCATAC 1251
DB |||||
QY 1905 ACGGCGCCCTCCAGGCCCCCTCGGGGCCATCCAGACCCCGTCTCACCCCTCTGCTCA 1964
DB |||||
QY 1252 ACGGCGCCCTCCAGGCCCCCTCGAGGCCATCCAGACCCCGCTCTCACCCCTCTGCTCA 1311
DB |||||

QY 1965 TGGTTCTCCAGCTCTGTCTCTACTATGCTTAAACGCAATGTCTTCAGAGGCGCTCCAAAG 2024
DB |||||
QY 1312 TGGTTCTCCAGCTCTGTCTCTACTATATGCTTAAACGCAATGTCTTCAGAGGACCCCCAAG 1371
DB |||||
QY 2025 GATGTCCCCAAAGGCCAGCGACATCTCTGAAATCACAGAGTTTCTGTGGGAGGGTTC 2084
DB |||||
QY 1372 GATGTCTCAAAAGGCACAGCGCCCTCGGAATCACAGAGTCTCTGTGGGAGGGTTC 1431
DB |||||
QY 2085 CATATCCAGTGGCTAGAAATTTGTATCCACAA CCCACCCAGTGAAGCAGCTACTCTCTCC 2144
DB |||||
QY 1432 CATGTCTAGTGGCTAGAAATTTGTATCCCAATCCCAAGTGAAGCAGCTGTCTCTCTCC 1491
DB |||||
QY 2145 AGTAGCAGGACCAAGTCCCTCGGGGGAAACGTGGTTCATCAGTGGTTCAGTGGGGTTCCAAAG 2204
DB |||||
QY 1492 AGTGGCAAGGACCAAGTCTCTGAGGGGGAAACGTGGTCTCAGTGGTTCAGTGGGGTTCCAAAG 1551
DB |||||
QY 2205 ATTATCCCTTAAACCTCATAGACCCAGGTCTCCACAGACAGAACAGTATTTGGAATAACCC 2264
DB |||||
QY 1552 GTTATCTCCAAACTCACAGACCCAGGTCTCCACAGGACAGCAGCAATTTGGAATACTCTCC 1611
DB |||||
QY 2265 CAGTGGGCGAGTTCTGTCTCTCCCAAGCTGGTATTTATTCCTCAACTGAAGCTGTGGCCAT 2324
DB |||||
QY 1612 CAGCGGGCTGTGTCTGTCTCTCCCAAGCTGGCATCATCTCCGAGAGGCGTTTCCAT 1671
DB |||||
QY 2325 GCCTATTCAGCTGCATCTCTTACGCTGCTAGTCTCTGCATCGAACAGAGCTGTACCCC 2384
DB |||||
QY 1672 GCCTGTCTCCGCGCATCTCCGACTCTCTGCCAGCCCTGCATCCAAACAGAGCACTGACCCC 1731
DB |||||
QY 2385 TTCTAGTAGGCTAAAGATTTCCAGGCTTCAAGATCAGAGGACAGAACTCTCTCTGCAGGGAA 2444
DB |||||
QY 1732 ATCTATTTAGGCAAAAGATTTCCAGGCTTCAAGATCAGAGGACAGAACTCTCTCTGCAGGGAG 1791
DB |||||
QY 2445 TAAAGAAATATTAACCCCAATGAAACATCAGTCTCTCAAGCTTCTCAAAAGCTGAAACAAAGG 2504
DB |||||
QY 1792 TAAAGAAATGTTAAAGCAAGTGAACATCACCTAGCTTTTCAAAAGCTGACAAACAAGG 1851
DB |||||
QY 2505 TATATCACAGTTGTTTCTGACATAGAAACAGATTTGATGATTTTAAAGAAATTTAAAGAA 2564
DB |||||
QY 1852 TATGTCCACAGTTGTTCTGAAACACAGAAACAGATTTGATGACTTAAAGAAAGTTTAAAGAA 1911
DB |||||
QY 2565 TGATTTTAGTTTACAGCCAAAGTTCTACTTCTCAATCTATGGATCAACTACTTAAACAAAA 2624
DB |||||
QY 1912 TGATTTTAGTTTACAGCCAAAGCTCTACATCTGAATCTATGGATCAACTACTAAGCAAAAA 1971
DB |||||
QY 2625 TAGAGAGGAGAAAAATCAAGAGATTTGATCAAAAGCAAAATTTGAACCAAGTGTAAAGGA 2684
DB |||||
QY 1972 TAGAGAGGAGAAAAAGTCAAGAGATTTGATTTAAAGATAAAACGGAAGCAAGTGTAAAGGA 2031
DB |||||
QY 2685 TTCCTTTCAATTGA-----AAATAGCAGCAGCAACTGTACCACTGGCAGCAGCAAGCCGAA 2738
DB |||||
QY 2032 TAGTTTTCATTGACAGCAGCAGCAGCAGCAACTGTACCACTGGCAGCAGCAAGACCAA 2091
DB |||||
QY 2739 TAGCCCCAGCATTTCCCTTCAATCTTAGTAACACGGAGCACAGAGGGGACCTCAGGT 2798
DB |||||
QY 2092 CAGCCCTAGACTCTCCCTTCCATGCTTAGTAAATGACAGAGCACAGAGGGGGCTCAGGT 2151
DB |||||
QY 2799 CACTTCCCAGGGGTTTCAGACTTCCAGCCAGCATGTAAACAGAGAAAGACGATAAAGGA 2858
DB |||||
QY 2152 CACATCCCAGGGGTGACAGACTTCCAGCCAGCCCTGCAACAAAGAGAGATGACAGAGA 2211
DB |||||
QY 2859 AGAGAAAGAGACGCGCTGAGCAAGTTAGGAAATCAACATTTGATCCCAATGCAAGGA 2918
DB |||||
QY 2212 AGAGAAAGAGACACAAACAGAGCAGGTTAGGAAATCGACATTTGAATTTCCCAATGCAAGGA 2271
DB |||||
QY 2919 GTTCAACCCAGCTTCTCTCAGCCAAAGCTTCTACTACCCCAACTTTCACCTCGGCC 2978
DB |||||
QY 2272 GTTCAACCCCTCGTCTTCTCTCAGCCAAAGCTTCTACTACCCCAACGTACCTCGGCC 2331
DB |||||
QY 2979 TCAAGCACAACTTAGCCCATCTATGTTGGGTTCATCAACAGCAACCTCCAGTTTATCTCA 3038
DB |||||
QY 2332 TCAAGCACAAACCCAGCCCATCTATGTTGGGTTCATCAGCAGCCAGCTCCAGTGTACACTCA 2391
DB |||||
QY 3039 GCCTGTTGTTTGTTCACCAAAATATGATGTATCCAGTCCCAGTGAGCCCGGTGCAACC 3098
DB |||||

Db 2392 GCTGTGTGCTTCGCACCAATATGATGATATCCCGTCCAGTGAAGCCGGCGGTACAC 2451
QY 3099 TTTATACCAATACCTATGACGCCCATGCCAGTGAATCAAGCAAGACATATAGAGC--- 3155
Db 2452 TTTATACCAATACCTATGACGCCCATGCCAGTGAATCAAGCAAGACATATAGAGCAGG 2511
QY 3156 ---AGTACCAATATGCCCAACAGCGGCAAGACAGCATCATCAGAGTGGCCATGATGCA 3212
Db 2512 TAAAGTACCAATATGCCCAACAGCGGCAAGACCAACATCATCAAGACCAATGATGCA 2571
QY 3213 CCCAGCGTCAGCAGCGGCGCCACCGATTCGACGCCACCCACCGCTTACTCCACGCAATA 3272
Db 2572 CCCAGCGTCCGGGAGGGCCACCCATCGTAGCCACCCCGCCGCTTACTCCACTCAGTA 2631
QY 3273 TGTGGCTTACAGTCTCAGCAGTTCGCCAAATCAGCGCCCTGTTCAGCATGTCACACATTA 3332
Db 2632 CGTTGGCTTACAGCCTCAGCAGTTCCCAATCAGCGCTTGTCCAGCATGTGCCCATTA 2691
QY 3333 TCAGTCTCAGCATCTCATGTCTATAGTCTGTATAGGTTAATACAGGGTAAATGCTAGAAATGATGC 3392
Db 2692 TCAGTCTCAGCATCTCATGTGTACAGTCTGTATACAGGTAAATGCCAGGATGATGCC 2751
QY 3393 ACCACCAACAGCGCCGAGCGCTGTAGTATCTTCACGCAACTCAGTACGGGGCTCA 3452
Db 2752 ACCACCAACAGCATGCTCAGCGCTGTAGTGTCTTCCTCAGCTGCTCAGTTCCGGGGCTCA 2811
QY 3453 TGAGCAGACGATGCGATGTATGATGTCCCAATTTACCATACACAGGAGACAAAGCCC 3512
Db 2812 CGAGCAGACGACGATGTATGATGTCCCAATTTACCATACACAGGAGACAAAGCCC 2871
QY 3513 TTTCTTTCTATTGTCATTTCCAGGGGTCCCTTGTCTCAGCAGTATGCGACCCCTAAAGC 3572
Db 2872 TTTCTTTCTATTGTCATTTCCAGGGGTCCCTTGTCTCAGCAGTATGCACTCTAATGC 2931
QY 3573 TACCTTGACCCACATATCTACACCCCTCAGCTTCAGTACCTACCCCATGTGACAGCAGCA 3632
Db 2932 CGCCCTGTATCCACATCTCCCATCTCCCTCAGCTTCGCGCCACTCCACCGACACAGCA 2991
QY 3633 AAGCCAAATGTTGAAGTCTATCTCGACCCAGTCTGTTCAGACCATCAGCACCAGGC 3692
Db 2992 AAGCAGCATGGTGAAGTCACTCTGACCCAGTCTGTTCAGACCATCAGCACCAGGC 3051
QY 3693 CGCCAGGCTCTCCATCTGGCCAGTCCACAGCAGTCAGCCATTTACCGCGGGGCT 3752
Db 3052 TGCCAGGCTCTTCATCTGGCCAGTCCACAGCAGTCAGCCATTTATCATGCGGGCT 3111
QY 3753 TGCGCCAACTCCACCTCCATGACACCTGCTCCACAGCGAGTCGCCACAGATAGTTT 3812
Db 3112 GGCACCAACACACCTTCCATGACACCTGCTCTAATACAGTCTCCACAGAGCAGTTT 3171
QY 3813 CCCAGCAGCAACAGAGTCTTTACGATCATCTCTCAGCTTCAGCCGCGGTATAC 3872
Db 3172 CCCAGCAGCAACAGAGTCTTTACCATCCATCCACCTTCTATGTTACGCCGATACAC 3231
QY 3873 CAACCCACCCACATGGGCCAGTACCTCAGGCTCATGTCAGTCAGGAATGGTTCCTTC 3932
Db 3232 CACCCACCCACATGGGCCAGTACCTCAGGCTCATGTCAGTCAGGAATGGTTCCTTC 3291
QY 3933 TCATCCAACTGCCCATGATGATGCTAATGACGACAGCCACCCGGGGTCCCA 3992
Db 3292 TCATCCAACTGCCCATGATGATGCTAATGACGACAGCCACCCGGGGTCCCA 3348
QY 3993 GGCGCCCTCGCTCAAGTGCATACAGCCCATTCAGTCTCGAACACAGCGCATTTCC 4052
Db 3349 GGCGCCCTCGCTCAAGTGCATACAGCCCATTCAGTCTCGAACACAGCGCATTTCC 3408
QY 4053 CTATATGACGACCTTTCAGTACAGCCCAACCAACAGCAGTTGTAAAGCTGCCCTGG 4112
Db 3409 TTATATGACGACCTTTCAGTACAGCCCAACCAACAGCAGTTGTAAAGCTGCCCTGG 3468
QY 4113 AGGAACCCAAAGGCCAAATTCCTCCTCTTACTGCTTCTACCACTGGAGACAG 4172

Db 3469 AGGAACCGAAAGGCCAAATCCCTTCTCCCTTCT-CTGCTTCTGCCAACCGAAGACACAG 3527
QY 4173 AAAAAGTAAATTTCAATTTATTTGTTTTTAAATATATATATGTTGTTTAAACATCC 4232
Db 3528 AAAAAGTAAATTTCAATTTGTTTTT-----TAAAGATACACTGATTTAAACATCT 3581
QY 4233 AATAGGAATGCTAAACAGTTCCTTGCAGTGAAGAT-ACITGGACCGAGTAGAGCATTT 4291
Db 3582 GATAGGAATGCTAAACAGTTCCTTGCAGTGAAGATGTTTTGGACCGAGTAGAGCATGT 3641
QY 4292 AGGAACCTTGGGGCTATTCCATAATTCATATGCTGTTTCAGATCCCGCAGGTACCCCA 4351
Db 3642 AGGACCTT-GTGGCTGTTCATAATTCATGCTGTTTGCAGGGTCTTGCAGTA-CCCA 3699
QY 4352 GCTCTGCTTGCAGAACTGGAAATTTATTTTAAATAACCTTGAAGTCTAATGAACA 4411
Db 3700 GCTCTGCTTGCAGAACTGGAAATTTATTTTAAATGAGGCTTGAAGTCTAATGAACA 3759
QY 4412 CATCAGCTAGCAAAAGAAAGTAAACAAGAGTGAATCTTGCT 4450
Db 3760 CATCAGCTAGCAAAAGAAAGTAAACAAGAGTGAATCTTGCT 3798

RESULT 11

AAF21628

ID AAF21628 standard; DNA; 2006 BP.

XX

AC AAF21628;

XX

DT 27-MAR-2001 (first entry)

XX

DE Human breast and ovarian cancer associated antigen gene SEQ ID 15.

XX

Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
nootropic; neuroprotective; antiviral; antiallergic; hepatotropic;
antidiabetic; antiinflammatory; antitumor; antileukemic; anticonvulsant;
antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
Addison's disease; allergy; autoimmune haemolytic anaemia;
autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
cardiovascular disorder; wound healing; neurological disease; ds.

XX

OS Homo sapiens.

XX

PN WO200055173-A1.

XX

PD 21-SEP-2000.

XX

PF 08-MAR-2000; 2000MO-US005881.

XX

PR 12-MAR-1999; 99US-0124270P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Ruben SM;

XX

DR WPI; 2000-611515/58.

XX

DR P-PSDB; AAB58725.

XX

New human breast and ovarian cancer associated gene sequences and the
polypeptides encoded by these genes, useful in the prevention, treatment
and diagnosis of cancer, immune disorders, cardiovascular disorders and
neurological diseases.

XX

PS Claim 1; Page 490-491; 1299pp; English.

XX

Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
proteins AAB58711 - AAB59128. The DNA and protein sequences are
associated with breast and ovarian cancer. Included in the invention are
sequences AAF22032 - AAF22040 and AAB59129 which are used in the
isolation and characterisation of the DNA and protein sequences of the
invention. The breast and ovarian cancer associated DNA, protein, agonist
or antagonist sequences exhibit cytostatic; immunosuppressive; nootropic;

CC

CC neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
CC antiinflammatory; antitumor; anticonvulsant; antibacterial;
CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
CC protein sequences are used in the diagnosis of cancer, particularly
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
CC and agonists may also be used in the diagnosis, prevention and treatment
CC of immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC cardiovascular disorders such as myocardial ischaemia; wound healing;
CC neurological diseases such as cerebral anoxia and epilepsy; and
XX infectious diseases

XX Sequence 2006 BP; 598 A; 577 C; 370 G; 456 T; 0 U; 5 Other;

Query Match 38.6%; Score 1730.6; DB 3; Length 2006;
Best Local Similarity 99.3%; Pred No. 0;
Matches 1768; Conservative 2; Mismatches 2; Indels 8; Gaps 3;

QY 2710 AACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCGAGATTTCCCTTCAATCTTAGT 2769
DB 6 AGCTGTACAGTGGCAGCAGCAAGCCGAATAGCCCGAGATTTCCCTTCAATCTTAGT 65

QY 2770 AACACGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACTCCAGCCCA 2829
DB 66 AACACGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACTTCCAGCCCA 125

QY 2830 GCATGTAAACAGAGAAAGACCATTAAGGAAGAGAAAGACGAGCTGAGCAAGTTAGG 2889
DB 126 GCATGTAAACAGAGAAAGACCATTAAGGAAGAGAAAGACGAGCTGAGCAAGTTAGG 185

QY 2890 AAATCAACATTGAATCCCAATCCAAAGGAGTTCAACCCACGTTCTCTCAGCCAAAG 2949
DB 186 AAATCAACATTGAATCCCAATCCAAAGGAGTTCAACCCACGTTCTCTCAGCCAAAG 245

QY 2950 CTTCTACTACCCCAACTTCACTGGGCTCAAGCAACACTAGGCCATCTATGGTGGT 3009
DB 246 CTTCTACTACCCCAACTTCACTGGGCTCAAGCAACACTAGGCCATCTATGGTGGT 305

QY 3010 CATCAACGCCAACTCCAGTTTATCTCAGCCGTTGTTTGGCCAAATATGATGAT 3069
DB 306 CATCAACGCCAACTCCAGTTTATCTCAGCCGTTGTTTGGCCAAATATGATGAT 365

QY 3070 CCAGTCCCAAGTCCAGCCGCGTCAACCTTTATACCCCAATCTATGACGCCATGCCA 3129
DB 366 CCAGTCCCAAGTCCAGCCGCGTCAACCTTTATACCCCAATCTATGACGCCATGCCA 425

QY 3130 GTGAATCAAGCCAAAGACATATAGAGC-----AGTACCAAAATATGCCCAAGCGGCAA 3183
DB 426 GTGAATCAAGCCAAAGACATATAGAGCAGGTAAGTACCAAAATATGCCCAAGCGGCAA 485

QY 3184 GACCCAGCATCAGAGTCCATGATGACCCAGCGTCCAGCGGGGCCACCGATTGCA 3243
DB 486 GACCCAGCATCAGAGTCCATGATGACCCAGCGTCCAGCGGGGCCACCGATTGCA 545

QY 3244 GCCACCCCAAGCTTACTCCAGCAATATGTTGCCCTACAGTCTCCAGCAGTTCCCAAT 3303
DB 546 GGNACCCCAAGCTTACTCCAGCAATATGTTGCCCTACAGTCTCCAGCAGTTCCCAAT 605

QY 3304 CAGCCCTTTGTTGAGATGTCGACATATCAGTCTCAGCATCTCATGTCTATAGTCT 3363
DB 606 CAGCCCTTTGTTGAGATGTCGACATATCAGTCTCAGCATCTCATGTCTATAGTCT 665

QY 3364 GTAAATACAGGTAACTGCTAGATGATGGCCACCAACACAGCCCGAGCTGTTTAGTA 3423
DB 666 GTAAATACAGGTAACTGCTAGATGATGGCCACCAACACAGCCCGAGCTGTTTAGTA 725

QY 3424 TCTTCTTCCAGCAACTCAGTACGGGGCTCATGACGACGATCGGATGATGATGCTCC 3483
DB 726 TCTTCTTCCAGCAACTCAGTACGGGGCTCATGACGACGATCGGATGATGATGCTCC 785

QY 3484 AAATTACCATACAAAGGAGCAAGCCCTTTCTTCTACTTTGCCATTTCCACGGGCTCC 3543

RESULT 12
ABV75289
ID ABV75289 standard; DNA; 1257 BP.
XX AC ABV75289;
XX

DB 786 AAATTACCATACAAAGGAGCAAGCCCTTCTTTCTACTTTTGCCATTTCCACGGGCTCC 845

QY 3544 CTTGCTCAGCAGTATGGC-ACCCTAAGCTTACCTGCACCCACATACTCCACACCTCTCA 3602

DB 846 CTTGCTCAGCAGTATGNGRACCTTAAGCTTACCTGTCACCCACATACTCCACACCTCTCA 905

QY 3603 GCCTTCAGCTACCCCTGAGCAGCAGCAAGCAACATGTTGGAAGTCACTCTGCAACC 3662

DB 906 GCCTTCAGCTACCCCTGAGCAGCAGCAAGCAACATGTTGGAAGTCACTCTGCAACC 965

QY 3663 CAGTCTCTGTTCCAGCACCATCAGCAGCAGCCGCCAGGCT-CTGCATCTGGCCAGTCCAC 3721

DB 966 CAGTCTCTGTTCCAGCACCATCAGCAGCAGCCGCCAGGCTTCTCATCTGGCCAGTCCAC 1025

QY 3722 AGCAGCAGTTCAGCTTTTACCAACCGGGGCTTTGGCCCAACTTCCACCTTCCATGACACCTG 3781

DB 1026 AGCAGCAGTTCAGCTTTTACCAACCGGGGCTTTGGCCCAACTTCCACCTTCCATGACACCTG 1085

QY 3782 CCTCCACACGCTGCGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTTACGA 3841

DB 1086 CCTCCACACGCTGCGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTTACGA 1145

QY 3842 TCCATCTCTTCTCAGCTTCAGCCGCGGTATACCAACCCACCCACATGCGCCACGCTACCTC 3901

DB 1146 TCCATCTCTTCTCAGCTTCAGCCGCGGTATACCAACCCACCCACATGCGCCACGCTACCTC 1205

QY 3902 AGGCTCATGTACAGTTCAGAAATGTTCTTCTCATCTCCAATGCCCCATGCGCAATGATGC 3961

DB 1206 AGGCTCATGTACAGTTCAGAAATGTTCTTCTCATCTCCAATGCCCCATGCGCAATGATGC 1265

QY 3962 TAATGACGACAGCAGCCCGCGGTCCCGAGCGCGCTCGCTCAAGAGTGCACGTACAGC 4021

DB 1266 TAATGACGACAGCAGCCCGCGGTCCCGAGCGCGCTCGCTCAAGAGTGCACGTACAGC 1325

QY 4022 CCATTCAGCTTCGACAAACAGCGCAATTTCCCTATATGACGACCCCTTCAGTACAAAGGCC 4081

DB 1326 CCATTCAGCTTCGACAAACAGCGCAATTTCCCTATATGACGACCCCTTCAGTACAAAGGCC 1385

QY 4082 ACCACCAACAGCAGTGTGAAGCTGCGCTGGAGGAACCGAAAGGCCAAATTCCTCTCTCC 4141

DB 1386 ACCACCAACAGCAGTGTGAAGCTGCGCTGGAGGAACCGAAAGGCCAAATTCCTCTCTCC 1445

QY 4142 CTTCTACTGCTTCTACCAACTGGAAGCAGCAGAAACTAGAAATTTCAATTTATTTTCTTTT 4201

DB 1446 CTTCTACTGCTTCTACCAACTGGAAGCAGCAGAAACTAGAAATTTCAATTTATTTTGT 1505

QY 4202 AAAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAATGCTAACAGTTCACCTGCACT 4261

DB 1506 AAAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAATGCTAACAGTTCACCTGCACT 1565

QY 4262 GGAAGATCTTGGACCGAGTAGAGGCATTTAGGAACTTTGGGGGCTATTCCATAATTTCCAT 4321

DB 1566 GGAAGATCTTGGACCGAGTAGAGGCATTTAGGAACTTTGGGGGCTATTCCATAATTTCCAT 1625

QY 4322 ATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGGCGAACTGGAAGTTATTTTA 4381

DB 1626 ATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGGCGAACTGGAAGTTATTTTA 1685

QY 4382 TTTTATTAATACCCCTTGAAGTTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAGAGTG 4441

DB 1686 TTTTATTAATACCCCTTGAAGTTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAGAGTG 1745

QY 4442 ATTTCTTGCTGCTATTACTGCTAAAAAATAAAAAAAAAAAAAA 4481

DB 1746 ATTTCTTGCTGCTATTACTGCTAAAAAATAAAAAAAAAAAAAA 1785

DT	07-MAR-2003	(first entry)
XX	Mouse SCA2 partial polypeptide encoding DNA.	
XX	SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity;	
XX	memory impairment; mutant; mouse; gene; db.	
KW		
KW		
XX	Mus musculus.	
OS		
XX		
XX	Key Location/Qualifiers	
PH	2..1255	
FT	/*tag= a	
FT	/product= "SCA2"	
FT	/partial	
FT	/note= "the start and stop codons are not indicated"	
XX		
XX	WO200289567-A2.	
PN		
XX		
XX	14-NOV-2002.	
PD		
XX		
XX	07-MAY-2002; 2002WO-US014698.	
PF		
PF		
XX		
XX	07-MAY-2001; 2001US-0289231P.	
PR		
XX		
XX	(CEDA-) CEDARS SINAI MEDICAL CENT.	
PA		
XX		
PI	Pulst SM;	
XX		
XX	WPI; 2003-103467/09.	
DR		
DR	P-PSDB; ABB82696.	
XX		
PT	New mutant mouse, useful as a disease model for identifying agents for	
PT	treating obesity or memory impairment, comprises a disrupted	
PT	spinocerebellar ataxia 2 gene.	
XX		
PS	Disclosure; Page 52-54; 76pp; English.	
XX		
CC	The invention relates to a mutant non-human mammal having a disrupted	
CC	spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a	
CC	disease model for identifying agents or compounds that may treat obesity	
CC	or memory impairment. The stem cell is used for generating the mutant	
CC	animal. The present sequence represents the partial nucleotide sequence	
CC	(GenBank Accession No. U70670) encoding a partial mouse SCA2 polypeptide	
XX		
SQ	Sequence 1257 BP; 312 A; 330 C; 358 G; 257 T; 0 U; 0 Other;	
Query Match 20.9%; Score 935.2; DB 10; Length 1257;		
Best Local Similarity 85.9%; Pred. No. 6.7e-171;		
Matches 1062; Conservative 0; Mismatches 168; Indels 6; Gaps 2;		
QY	685 CAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCGCGCGCGCGCGCGCAAT 744	
Db	2 CACAGGGCGCGCTACCATGTCTGAAGCGGAGCGCGAGCGCGCGCGCGCGCGCACT 61	
QY	745 GTCGCGAAGCCCGCGCGCAGCGCGCTTCTAGCGTGCCTCGCGCGCGCGCGCGCGTCC 804	
Db	62 GGCGCGAAGCCCGCGCGCGCGCTCTCTCGTGCCTCGCGCGCGCGCGCGCGCGCG 121	
QY	805 TCGTCTCTGGTCTCTGTCTCGGCGACAGCGCTCCTCTCTCGGTG---GTGCGGCGGAC 861	
Db	122 GTGACCTTCGGGTTCGCTGTGTGTGTCGCGCGCGCGCGCGGTGGCGTCTTCTCGGCGG 181	
QY	862 TCCGGCGCGGAGCGCGCGCTCGGCAGAGGTGMAACAGTAAACAGGACTGCCTCAG 921	
Db	182 GCGGCGCGCGCGGTCCCGCGCTGGGCAGAGGTGGGAACAGTAGCAAGGACTGCCTCAG 241	
QY	922 TCTACGATTTCTTTTGTGGAATCTATGCAAAATATGAGGATGGTTTCATATCTTACATCA 981	
Db	242 CCTACGATTTCTTTTGTGGAATCTATGCAAAACGTGAGGATGGTTTCATATCTTACGTC 301	
QY	982 GTTGTTGGCTCCAAATGTGAAGTACAAGTGAAGAAATGAGGATATATATGAAGAGATTTT 1041	
Db	302 GTTGTTGGATCGAAATGTGAAGTACAAGTGAAGAAACGAGGATATATGAAGAGATTTT 361	

KW osteoporosis; neurological disease; Alzheimer's disease;
 KW Parkinson's disease; dementia; short memory; cancer;
 KW sense or motor function; emotional reaction; fear response; panic;
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
 KW tranquiliser.
 XX
 OS Homo sapiens.
 XX
 PN EP1447413-A2.
 XX
 XX 18-AUG-2004.
 XX
 PF 12-FEB-2004; 2004EP-00003145.
 XX
 XX 14-FEB-2003; 2003JP-00102207.
 PR 09-MAY-2003; 2003JP-00131452.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;
 XX
 DR WPI: 2004-583265/57.
 DR P-PSDB; ADR10292.
 XX
 PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 PS Claim 1; SEQ ID NO 1842; 2686pp; English.
 CC This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNAs. The present invention
 CC describes an immunoassay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
 CC cytostatic and tranquiliser activities. This polynucleotide is a full
 CC length human cDNA sequence of the invention. NOTE: This sequence is not
 CC given in the sequence listing of the specification but can be obtained on
 CC CD-ROM from the European Patent Office, Vienna Sub-office.
 XX
 SQ Sequence 4369 BP; 1047 A; 1122 C; 1051 G; 1149 T; 0 U; 0 Other;
 Query Match 14.1%; Score 633; DB 13; Length 4369;
 Best Local Similarity 92.8%; Pred. No. 2.7e-112;
 Matches 697; Conservative 0; Mismatches 0; Indels 54; Gaps 1;
 QY 3154 GCAGTACCAATATGCCCAACAGCGGCAAGACAGCATCATCAGAGTGCATGATGCAC 3213
 DB 96 GCAGTACCAATATGCCCAACAGCGGCAAGACAGCATCATCAGAGTGCATGATGCAC 155
 QY 3214 CCAGCGTCAGAGGGGGCCCGGATTCAGCCACCCAGCTTACTCCAGCAATAT 3273
 DB 156 CCAGCGTCAGAGGGGGCCCGGATTCAGCCACCCAGCTTACTCCAGCAATAT 215
 QY 3274 GTTGCTTACAGTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATAT 3333
 DB 216 GTTGCTTACAGTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATAT 275
 QY 3334 CAGTCTCAGCATCCTCATGTCTATAGTCCTGTAATACAGGGTAATGCTAGAAATGATGGCA 3393
 DB 276 CAGTCTCAGCATCCTCATGTCTATAGTCCTGTAATACAGGGTAATGCTAGAAATGATGGCA 335
 QY 3394 CCACCAACACAGCGCGGCTGGTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCAT 3453
 XX
 DB 336 CCACCAACACAGCGCGGCTGGTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCAT 395
 QY 3454 GAGCAGAGCGATGCGATGTATGTCATGTGCCAAATTTACCATACAAAGAGACAAGCCCT 3513
 DB 396 GAGCAGAGCGATGCGATGTATG- 417
 QY 3514 TCTTTCTACTTTGGCCATTTTCCAGCGGGCTCCCTTGTGCTCAGCAGTATGCGCACCTTAACGCT 3573
 DB 418 -----TTTCCACGGGGCTCCCTTGTGCTCAGCAGTATGCGCACCTTAACGCT 461
 QY 3574 ACCCTGCAACCCACATATCTCCACACCCCTCAGCCTTCAGCTTACCCCTTGGACAGAGCAA 3633
 DB 462 ACCCTGCAACCCACATATCTCCACACCCCTCAGCCTTCAGCTTACCCCTTGGACAGAGCAA 521
 QY 3634 AGCCAAACATGTGGGAAGTCACTCTGCACCCAGTCTCTGTTTCAGCAGCATCAGCACAGGGCC 3693
 DB 522 AGCCAAACATGTGGGAAGTCACTCTGCACCCAGTCTCTGTTTCAGCAGCATCAGCACAGGGCC 581
 QY 3694 GCCCAGGGCTCTCCATCTGSCCAGTCCACAGCAGCAGTCCAGCCATTTACCACCGGGGCTT 3753
 DB 582 GCCCAGGGCTCTCCATCTGSCCAGTCCACAGCAGCAGTCCAGCCATTTACCACCGGGGCTT 641
 QY 3754 GCGCCAACTCCACCCCTCCATGACACCTGCTCCAAACAGCAGTCCGCCACAGAAATAGTTTC 3813
 DB 642 GCGCCAACTCCACCCCTCCATGACACCTGCTCCAAACAGCAGTCCGCCACAGAAATAGTTTC 701
 QY 3814 CCAGCAGCAACACAGCAGTCTTTTACGATCCATCTCTCAGCTTCAGCGCGGGTATACC 3873
 DB 702 CCAGCAGCAACACAGCAGTCTTTTACGATCCATCTCTCAGCTTCAGCGCGGGTATACC 761
 QY 3874 AACCCACCCACATGGCCACCTTACTCAGG 3904
 DB 762 AACCCACCCACATGGCCACCTTACTCAGG 792
 RESULT 14
 AAV17229
 ID AAV17229 standard; DNA; 623 BP.
 AC AAV17229;
 DT 29-JUN-1998 (first entry)
 XX SCA2 gene fragment.
 DE SCA2 gene; spinocerebellar ataxis type II; CAG repeat; PCR primer; ss.
 XX Synthetic.
 OS
 FH Key Location/Qualifiers
 CDS 341..583
 FT /*tag= a
 FT /note= "SCA2 protein fragment, no stop codon given"
 XX
 PN WO9803679-A1.
 XX
 XX 29-JAN-1998.
 PD
 XX
 PF 18-JUL-1996; 96WO-JP001999.
 XX
 XX 18-JUL-1996; 96WO-JP001999.
 PR
 XX
 PA (SRLS-) SRL INC.
 XX
 PI Teuji S, Sanpei K;
 XX
 XX WPI; 1998-120796/11.
 DR P-PSDB; AAW41372.
 XX
 XX Diagnosing spinocerebellar ataxis type II - by PCR and determining number
 PT of CAG repeat units.
 XX
 XX Example 1; Page 11-12; 23pp; Japanese.
 PS

PD	13-NOV-1997.
XX	
PF	08-MAY-1997; 97WO-US007725.
XX	
PR	08-MAY-1996; 96US-0017388P.
PR	19-JUL-1996; 96US-0022207P.
PR	08-OCT-1996; 96US-0072708A.
XX	(CEDA-) CEDARS SINAI MEDICAL CENT.
PA	Pulst S;
PI	
XX	
DR	WFI; 1998-086523/08.
XX	
PT	Nucleic acids encoding human and mouse ataxin 2 - a product of the
PT	spinocerebellar ataxia 2 gene, SCA2; useful in the diagnosis of ataxia
PT	type 2.
XX	
PS	Example 2; Page 51-52; 98pp; English.
XX	
CC	This genomic DNA in plasmid PL65I22B includes a CAG repeat region from
CC	the novel human SCA2 gene (see AAV06552). It was identified following the
CC	construction of a bacterial artificial chromosome contig and a P1
CC	artificial chromosome of the spinocerebellar ataxia 2 (SCA2) gene region
CC	and the identification of the SCA2 gene from this contiguous map unit
CC	using a technique that screens for the presence of DNA trinucleotide
CC	repeats. The SCA2 locus is at 12q24.1. Ataxia type 2 can be diagnosed by
CC	detecting a genomic or transcribed mRNA sequence in an individual having
CC	an expanded CAG repeat at a location corresponding to the CAG repeat
CC	region of the SCA2 gene. The presence of at least 13 CAG repeats above
CC	the normal level (22, occasionally 23, repeats) is indicative of SCA2.
CC	Primers (see AA99640-41) amplifying at least this region are used for
CC	diagnosis. Also claimed are full-length ataxin-2 cDNAs for human and
CC	mouse (see AAV06552-53), kits for detecting mutations at the SCA2 locus,
CC	antisense oligonucleotides, and transgenic animals useful for studying
CC	the physiological roles of SCA2 polypeptide (ataxin-2, see AAW33807-08)
CC	and its effect upon behaviour
XX	
SQ	Sequence 516 BP; 50 A; 228 C; 166 G; 72 T; 0 U; 0 Other;
	Query Match 11.1%; Score 497.4; DB 2; Length 516;
	Best Local Similarity 99.8%; Pred. No. 2.6e-86;
	Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0
Qy	392 TTGGTAGCAGCGCAAACGCGCGCGCGGTTCGGGCCCGGCTCCGCGCGGCTCCTTGCGTC 451
Dd	
	1 TTGGTAGCACGGAACGCGCGCGCGGCTTCGGGCCCGGCTCCGCGCGGCTCCTTGCGTC 60
Qy	452 TCGGCGGGCGCTCCCGGCCCTTCGTCGTCGTCCTTCTCCCCCTGCACAGCCGGCGCCC 511
Dd	
	61 TCGGCGGGCGCTCCCGGCCCTTCGTCGTCGTCCTTCTCCCCCTGCACAGCCGGCGCCC 120
Qy	512 CTCGGCGCGCGCAACCGCGCGCTCCCGCTCGGGGCCCGCTGCTGCCCGCGCGTTCCG 571
Dd	
	121 CTCGGCGCGCGCAACCGCGCGCTCCCGCTCGGGGCCCGCTGCTGCCCGCGCGTTCCG 180
Qy	572 GCGTCTCTTCGGCGCGCGCGCTCCCGGCTGFCGCCGCCCGCGCTGCGAGCGGCTGATG 631
Dd	
	181 GCGTCTCTTCGGCGCGCGCGCTCCCGGCTGTCGCCGCCCGCGCTGCGAGCGGCTGATG 240
Qy	632 GGCCCTCACCATTGCTCGCTGAAGCCCCACAGCAGCAGCAGCAGCAGCAACAGCAGC 691
Dd	
	241 GGCCCTCACCATTGCTCGCTGAAGCCCCACAGCAGCAGCAGCAGCAGCAACAGCAGC 300
Qy	692 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGTCCGGA 751
Dd	
	301 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGTCCGCA 360
Qy	752 AGCCCGGGGAGCGGCGCTTCTAGCGTCCCGCGCGCGCGCTTTCGCGCTCCTCGTCT 811
Dd	
	361 AGCCCGGGGAGCGGCGCTTCTAGCGTCCCGCGCGCGCGCTTTCGCGCTCCTCGTCT 420
Qy	812 CGGTCTCTCGTTCGTCGCGCACGGCTCCCTCTCTCTCGGTGCTCGCGCGACCTCCGCGCGCG 871

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 17:46:01 ; Search time 15479 Seconds
(without alignments)
13544.345 Million cell updates/sec

Title: US-10-802-228-1
Perfect score: 4481
Sequence: 1 acccccggagaaagcaaccca.....taaaaaaaaaaaaaaaaaaaaaa 4481

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1327	29.6	1528	4	CR619107 full-leng
2	903	20.2	912	5	BQ883680 AGENCOURT
3	899.8	20.1	1000	5	EX402110
4	899	20.1	1060	1	AL554896
5	897	20.0	938	5	EX375946
6	850.6	19.0	977	7	CN646951 ILLUMIGEN
7	842	18.8	1100	3	BM455214 AGENCOURT
8	822.8	18.4	859	5	BQ716397
9	819.2	18.3	889	6	CA489164
10	797.6	17.8	846	1	AU124606
11	785.2	17.5	1066	1	AL554863
12	773.6	17.3	863	6	CD616049
13	767	17.1	830	6	CD616048
14	761.2	17.0	877	7	CO886647
15	753.2	16.8	784	6	CA775370
16	747	16.7	750	-1	AU142284
17	744.6	16.6	940	5	BUI59977
18	739.6	16.5	877	5	BQ428363
19	736.6	16.4	856	7	CJ039644
20	736	16.4	878	6	CD106563
21	732.6	16.3	779	1	AU121620
22	729.8	16.3	804	2	BG201697 RST21039

ALIGNMENTS

RESULT 1		CR619107		1528 bp mRNA linear HTC 21-JUL-2004	
LOCUS		full-length cDNA clone CS0D1087Y117 of Placenta Cot 25-normalized		of Homo sapiens (human)	
DEFINITION		CR619107		CR619107.1 GI:50499914	
ACCESSION		CR619107		HTC; CNSLT cDNA	
VERSION		CR619107		Homo sapiens (human)	
KEYWORDS		Homo sapiens		Homo sapiens	
SOURCE		Homo sapiens		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1 (bases 1 to 1528)		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
AUTHORS		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.		Full-length cDNA libraries and normalization	
TITLE		Unpublished		Contact : Feng Liang Email : fliang@lifetech.com URL :	
JOURNAL		http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		Faraday Avenue	
REMARK		2 (bases 1 to 1528)		Genoscope.	
REFERENCE		Direct Submission		Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :	
AUTHORS		BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr		- Web : www.genoscope.cns.fr)	
TITLE		1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime		into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library	
JOURNAL		was normalized. Library was constructed by Life Technologies, a		division of Invitrogen.	
REMARK		Location/Qualifiers		1..1528	
FEATURES		/organism="Homo sapiens"		/mol_type="mRNA"	
source		/db_xref="taxon:9606"		/clone="CS0D1087Y117"	
ORIGIN		/tissue_type="Placenta Cot 25-normalized"		/plasmid="pCMVSPORT_6"	
Query Match		29.6%		Score 1327; DB 4; Length 1528;	
Best Local Similarity		89.7%		Pred. No. 3.5e-292;	
Matches 1522; Conservative		0; Mismatches		175; Gaps	
		0; Indels		2;	

QY	2743	CCACGATTTCCCTTCAATACTTAGTAAACGAGGACAAAGAGGGGACTGTGAGGTCACT	2802
DB	1	CCACGATTTCCCTTCAATACTTAGTAAACGAGGACAAAGAGGGGACTGTGAGGTCACT	60
QY	2803	TCCCAGGGGTTACAGCTTCAGGCCACGATGTAAACAAGAGAAGACGATNAGGAAG	2862
DB	61	TCCCAGGGGTTACAGCTTCAGGCCACGATGTAAACAAGAGAAGACGATNAGGAAG	120
QY	2863	AAGAAAGACGCAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAAGGAGTTC	2922
DB	121	AAGAAAGACGCAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCAAAGGAGTTC	180
QY	2923	AAACCAAGTTCCTTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACCTCGGCCCTCAA	2982
DB	181	AAACCAAGTTCCTTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACCTCGGCCCTCAA	240
QY	2983	GCACAACCTAGCCCATCTATGGTGGGTCAACAGGCCAACTCAGTTTATATCTCAGCCT	3042
DB	241	GCACAACCTAGCCCATCTATGGTGGGTCAACAGGCCAACTCAGTTTATATCTCAGCCT	3000
QY	3043	GTTCGTGTTTGACCAAAATATGATGTATCCAGTCCAGTGAGCCAGCGCTGCACACCTTTA	3102
DB	301	GTTCGTGTTTGACCAAAATATGATGTATCCAGTCCAGTGAGCCAGCGCTGCACACCTTTA	360
QY	3103	TACCAATACTTATGACGCCCATGCCAGTGAATCAAGCCAAAGACATATAGAGC-----A	3156
DB	361	TACCAATACTTATGACGCCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGGTAAA	420
QY	3157	GTACCAATATGCCCAACAGGGGCAAGACGAGCATCATCAGAGTGCGATGTCACCCA	3216
DB	421	GTACCAATATGCCCAACAGGGGCAAGACGAGCATCATCAGAGTGCGATGTCACCCA	480
QY	3217	GGGTACAGAGGGGCCACCGATTGACGCCACCCACCGCTTACTCCAGCGCAATATGTT	3276
DB	481	GGGTACAGAGGGGCCACCGATTGACGCCACCCACCGCTTACTCCAGCGCAATATGTT	540
QY	3277	GCCTACAGTCTCAGCAGTTCCCAATCAGGCCCTTGTTACGACATGTGCCACATTTACG	3336
DB	541	GCCTACAGTCTCAGCAGTTCCCAATCAGGCCCTTGTTACGACATGTGCCACATTTACG	600
QY	3337	TCTCAGATCTCATGTCTTAGTCTGTATATACAGGTAATGCTAGATGATGGCAACA	3396
DB	601	TCTCAGATCTCATGTCTTAGTCTGTATATACAGGTAATGCTAGATGATGGCAACA	660
QY	3397	CCAAACAACGCCACGCTGGTTTACTATCTTCTTCAGCAACTCAGTACGGGCTCATGAG	3456
DB	661	CCAAACAACGCCACGCTGGTTTACTATCTTCTTCAGCAACTCAGTACGGGCTCATGAG	720
QY	3457	CAGACGATGCGATGTATGCATGTCCAAATATACATAACAACAGGAGACAAGCCCTTCT	3516
DB	721	CAGACGATGCGATGTATGCATGTCCAAATATACATAACAACAGGAGACAAGCCCTTCT	780
QY	3517	TTCTACTTTGCCATTTCAACGGGCTCCCTGCTCAGCAGTATGCGCACCTTAACGCTACC	3576
DB	781	TTCTACTTTGCCATTTCAACGGGCTCCCTGCTCAGCAGTATGCGCACCTTAACGCTACC	840
QY	3577	CTGCACCCACATCTCCACAACCTCAGCCTTACGTAACCCCACTGCACACGACGACAAAGC	3636
DB	841	CTGCACCCACATCTCCACAACCTCAGCCTTACGTAACCCCACTGCACACGACGACAAAGC	900
QY	3637	CAACAATGTTGGAAGTCATCTCTGCACCCAGTCTGTTTACGACACCATCAGCACCGCGCC	3696
DB	901	CAACAATGTTGGAAGTCATCTCTGCACCCAGTCTGTTTACGACACCATCAGCACCGCGCC	960
QY	3697	CAGGCTTCTCATCTGGCCAGTGCACAGACGAGTGCAGCATTTTACACGCGGGCTTGGC	3756
DB	961	CAGGCTTCTCATCTGGCCAGTGCACAGACGAGTGCAGCATTTTACACGCGGGCTTGGC	1020
QY	3757	CCAACTCCACCTCCATCAGCACTGCTCCAAACGCGAGTGCACACAGATAGTTTCCCA	3816
DB	1021	CCAACTCCACCTCCATCAGCACTGCTCCAAACGCGAGTGCACACAGATAGTTTCCCA	1080
QY	3817	GCAGCAACAACAGCTGTCTTTACGATTCATCTTCTCAGTTCAGCGGGGTATACCAAC	3876

[illegible]

```
source
1. .912
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6179729"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupeki dorsal root ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TGACCCAGCGGCC-3' and
5'-GACTAGTTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupeki, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN
Query Match 20.2%; Score 903; DB 5; Length 912;
Best Local Similarity 99.3%; Pred. No. 3.4e-195;
Matches 906; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2281 GCTTCTCCCAAGCTGGTATTATTCACACTGAAGCTGTGGCCATGCTATTCCAGCTGCA 2340
DB 1 GCTTCTCCCAAGCTGGTATTATTCACACTGAAGCTGTGGCCATGCTATTCCAGCTGCA 60
QY 2341 TCTCTACGCTCTAGTCTGATCGAACAGAGCTGTACCCCTCTAGTGAGGCTAA 2400
DB 61 TCTCTACGCTCTAGTCTGATCGAACAGAGCTGTACCCCTCTAGTGAGGCTAA 120
QY 2401 GATTTCAGGCTTCAAGATCAGAGCGAGAAGCTCTCTGAGGGAATAAGAAATATTAAA 2460
DB 121 GATTTCAGGCTTCAAGATCAGAGCGAGAAGCTCTCTGAGGGAATAAGAAATATTAAA 180
QY 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAAGTTGT 2520
DB 181 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAAGTTGT 240
QY 2521 TCTGAACATGAAGAACAGATTGATGATTTAAGAAATTTAAGATGATTTAGTTACAG 2580
DB 241 TCTGAACATGAAGAACAGATTGATGATTTAAGAAATTTAAGATGATTTAGTTACAG 300
QY 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAATAAGAGGAGGAGAAAA 2640
DB 301 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAATAAGAGGAGGAGAAAA 360
QY 2641 TCAAGAGATTGTGATCAAGAGAAAAATTGAACCAAGTCTAAGGATTTCTTTTATTGAAAT 2700
DB 361 TCAAGAGATTGTGATCAAGAGAAAAATTGAACCAAGTCTAAGGATTTCTTTTATTGAAAT 420
QY 2701 AGCAGAGCAACTGTGTACAGTGGCAGCAGCAGCGCAATAGCCGACATTTCCCTTCA 2760
DB 421 AGCAGAGCAACTGTGTACAGTGGCAGCAGCAGCGCAATAGCCGACATTTCCCTTCA 480
QY 2761 ATACTTAGTAACAGGAGCAGAGAGGAGCTGAGTCACTTCCCAAGGGGTTTCAGACT 2820
DB 481 ATACTTAGTAACAGGAGCAGAGAGGAGCTGAGTCACTTCCCAAGGGGTTTCAGACT 540
QY 2821 TCCAGCCCGCAGCATGTAAACAGAGAAAAAGCAGTAAGGAAGAGAAAGACGCGCTGAG 2880
DB 541 TCCAGCCCGCAGCATGTAAACAGAGAAAAAGCAGTAAGGAAGAGAAAGACGCGCTGAG 600
QY 2881 CAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGATTCAACCCAGCTTCTCTCT 2940
DB 601 CAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGATTCAACCCAGCTTCTCTCT 660
QY 2941 CAGCCAAAGCTTCTACTACCCCAACTTCACTCGGCTCAAGCAGCAACCTAGCCCATCT 3000
DB 661 CAGCCAAAGCTTCTACTACCCCAACTTCACTCGGCTCAAGCAGCAACCTAGCCCATCT 720

1000 bp mRNA linear EST 01-MAY-2004
BX402110 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC015YB14 5-PRIME, mRNA sequence.
BX402110
BX402110.2 GI:46922451
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1000)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30632041.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6451.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS1AC0042F080Plk=6451.r.
Location/Qualifiers
1. .1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC015YB14"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
1. .1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC015YB14"
/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 20.1%; Score 899.8; DB 5; Length 1000;
Best Local Similarity 96.7%; Pred. No. 1.9e-194;
Matches 927; Conservative 10; Mismatches 18; Indels 4; Gaps 2;

QY 327 TCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 386
DB 1 TCCCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
QY 387 CTGTTTGTAGTACAAAGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 446
DB 61 CTGTTTGTAGTACAAAGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
```

```
QY 447 TGGTCTCGCGGGCTCCCGCCCTTCTGTCGTCGTCCTTCTCCCTCGCGAGCCCGGG 506
Db |||
QY 507 CGCCCTCGCGCGCCCAACCGGCTCCCGCTCGCGGCGCGTGGTCCCGCGCGG 566
Db |||
QY 567 TTCGCGGCTCTCTTGGCGCGCCCGCTCCCGCTGTCGCGCGCGGCTGGAGCCCGT 626
Db |||
QY 627 GTATGGGCGCTCACTGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAACA 686
Db |||
QY 687 GCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACA 743
Db |||
QY 744 TGTCCGCAAGCCCGCGCGCAGCGG-CCTTCTAGCGTCGCCCGCGCGCTTCGCCGT 802
Db |||
QY 803 CCTGCTCTCGGTCTCCTCGCTCGGCGCAAGGCTCCCTCTCGGTGTCGCGGCGACCT 862
Db |||
QY 863 CCGCGCGGGAGCGCCCGCTGGCAGAGGTGCAACAGTAAAGAGCTGCTCAGT 922
Db |||
QY 923 CTACGATTTCTTTTGATGGAATCTATGCAAAATATGAGATGTTTCAATATCTTACATCAG 982
Db |||
QY 983 TTGTTGGCTCCAAATGTGAAGTACAAAGTCAAAATGGAGGTATATATGAAGAGTTT 1042
Db |||
QY 1043 AAATTACAGTCCGAAGTGTGATTTGGTACTGTATGTCGCGCATGAGAAAGTACAGAT 1102
Db |||
QY 1103 CCAGTTCGCGGCGGAAACGCTGAAGAAATAATGGAGAGTATTTGTTCAAAATGTTCCAGACT 1162
Db |||
QY 1163 TTGTTGGTACAGTTTAAAGATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTG 1222
Db |||
QY 1223 ACTCTGCTATCAGTCTAAAGTGAATGCGGACCAACAAAGAGAGACCTGGAGCCCTGG 1281
Db |||
QY 901 ACTCTGCTATCAGTCTAAAGTGAATGCGGAAACCAAAAGAGAGAGACCTGGAGCCCTGGG 959

RESULT 4
AL554896 1060 bp mRNA linear EST 30-WAR-2004
LOCUS AL554896 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI087Y117 5-PRIME, mRNA sequence.
ACCESSION AL554896
VERSION AL554896.3 GI:45859645
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1060)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
```

```
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31276706.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6451.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DI087AE09QPI&c=6451.r.
FEATURES
source
1..1060
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI087Y117"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-Oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 20.1%; Score 999; DB 1; Length 1060;
Best Local Similarity 97.3%; Pred. No. 2.9e-194;
Matches 941; Conservative 6; Mismatches 12; Indels 8; Gaps 3;
QY 2743 CCCAGCATTTCCCTTCAATCTTAGTAACAGCGAGCAGAGGGGACCTGAGGTCACT 2802
Db 1 CCCAGCATTTCCCTTCAATCTTAGTAACAGCGAGCAGAGGGGACCTGAGGTCACT 60
QY 2803 TCCCAAGGGGTTTCCAGACTTCCAGCCCGAGCATGTAACAAGAGAAAGAGCATAAAGAAAG 2862
Db 61 TCCCAAGGGGTTTCCAGACTTCCAGCCCGAGCATGTAACAAGAGAAAGAGCATAAAGAAAG 120
QY 2863 AAGAAAGAGCGAGTGTAGGAAATCAACATTGGAATCCCAATGCAAGAGAGTTC 2922
Db 121 AAGAAAGAGCGAGTGTAGGAAATCAACATTGGAATCCCAATGCAAGAGAGTTC 180
QY 2923 AACCCAGTTCCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAA 2982
Db 181 AACCCAGTTCCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAA 240
QY 2983 GCACAACTTAGCCCATCTATGGTGGGTGTCATCAACAGCCCAACTCCAGCTTTTATCTCAGGCT 3042
Db 241 GCACAACTTAGCCCATCTATGGTGGGTGTCATCAACAGCCCAACTCCAGCTTTTATCTCAGGCT 300
QY 3043 GTTTGTTTGGACCAAAATATGATGTATCCAGTCCAGTGAGCCGAGGCGTCAACCTTTTA 3102
Db 301 GTTTGTTTGGACCAAAATATGATGTATCCAGTCCAGTGAGCCGAGGCGTCAACCTTTTA 360
QY 3103 TACCAATACCTATGAGCCCATGCGAGTGAATCAAGCAAGACATATAGAGC-----A 3156
Db 361 TACCAATACCTATGAGCCCATGCGAGTGAATCAAGCAAGACATATAGAGAGAGTAAA 420
QY 3157 GTACCAAAATATGCCCAACAGCGCAAGACCATCATCAGAGTGCCATGATGACCCCA 3216
Db 421 GTACCAAAATATGCCCAACAGCGCAAGACCATCATCAGAGTGCCATGATGACCCCA 480
QY 3217 GCGTCAGAGCGGCGCCACCGATTGTCAGCCACCCCACTTACTCCACCAATATGTT 3276
Db 481 GCGTCA-CAGGCGGCCCAACCGATTGTCAGCCACCCCACTTACTCCACCAATATGTT 539
QY 3277 GCCTTACAGTCTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATTCAG 3336
Db 540 GCCTTACAGTCTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATTCAG 599
```

```
QY 3337 TCTCAGCATCCTCATGCTCTATAGTCTCTGTAATACAGGGTAATGCTAGAAATGATGGCACCA 3396
Db 600 TCTCAGCATCCTCATGCTCTATAGTCTCTGTAATACAGGGTAATGCTAGAAATGATGGCACCA 659
QY 3397 CCAACACACGCCAGCCTGGTTTAGTATCTTCTTACAGCAACTCAGTACCGGGCTCATAGAG 3456
Db 660 CCAACACACGCCAGCCTGGTTTAGTATCTTCTTACAGCAACTCAGTACCGGGCTCATAGAG 719
QY 3457 CAGACGCATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3516
Db 720 CAGACGCATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
QY 3517 TTCTACTTTGGCCATTTCCAGGGCTCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACC 3576
Db 780 TTCTACTTTGGCCATTTCCAGGGCTCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACC 839
QY 3577 CTGACCCACATATCTCCACACCTCAGCCTTCTCAGCTACCCCACTCGACAGCAGCAAGC 3636
Db 840 CTGACCCACATATCTCCACACCTCAGCCTTCTCAGCTACCCCACTCGACAGCAGCAGC-AARC 898
QY 3637 CAACATGTTGGAGTATCTCTGACCCAGTCTCTGTTTTCAGCACCATCAGCACCGGCGCC 3696
Db 899 CAACATGTTGGAGTATCTCTGACCCAGTCTCTGTTTTCAGCACCATCAGCACCGGCGCC 958
QY 3697 CAGGCTC 3703
Db 959 AGCTCTC 965

RESULT 5
BX375946
LOCUS
DEFINITION
BX375946 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC020YG10 5-PRIME, mRNA sequence.
ACCESSION
BX375946
VERSION
BX375946.2 GI:46573234
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 938)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 8, 2003 this sequence version replaced gi:3048458.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6451.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DC020BD05QP1&c=6451.r.
Location/Qualifiers
1. 938
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC020YG10"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source

ORIGIN
CN646951 977 bp mRNA linear EST 13-MAY-2004
LOCUS
DEFINITION ILLUMIGEN_MCO_27453 Katze_MBR Macaca mulatta cDNA clone IB1UW:8082
```

```
Query Match 20.0%; Score 897; DB 5; Length 938;
Best Local Similarity 98.6%; Pred. No. 8e-194;
Matches 919; Conservative 3; Mismatches 8; Indels 2; Gaps 2;
QY 2228 CCAGGCTCTCCACAGACAGACAGTATTTGAAATACCCAGTGGGCGAGTCTTCTGCTTCTC 2287
Db 1 CCCGGATCCACAGACANACAGTATTTGAAATACCCAGTGGGCGAGTCTTCTGCTTCTC 60
QY 2288 CCCAAGCTGGTATTTATTCACACTGAAGCTGTTGCCATGCTATTTCCAGCTGCATCTCTTA 2347
Db 61 CCCAAGCTGGTATTTATTCACACTGAAGCTGTTGCCATGCTATTTCCAGCTGCATCTCTTA 120
QY 2348 CGCCTGCTAGTCTCGATCGACAGAGCTGTTTACCCCTTCTAGTGAGGCTAAAGATTCCA 2407
Db 121 CGCCTGCTAGTCTCGATCGACAGAGCTGTTTACCCCTTCTAGTGAGGCTAAAGATTCCA 180
QY 2408 GCGCTTCAAGATCAGAGGAGCAAACTCTCTCGCAGGGAATAAGAAAAATATTAAACCCAATG 2467
Db 181 GCGCTTCAAGATCAGAGGAGCAAACTCTCTCGCAGGGAATAAGAAAAATATTAAACCCAATG 240
QY 2468 AAACATCACTAGCTTCTCAAAAGCTGAAAAAAGGATATATCACCAGTGTGTTTCTGAAC 2527
Db 241 AAACATCACTAGCTTCTCAAAAGCTGAAAAAAGGATATATCACCAGTGTGTTTCTGAAC 300
QY 2528 ATAGAAAAAGC-ATTGATGATTTAAAGAAATTTAAGAATGATTTTAGGTTTACAGCCAAGT 2586
Db 301 ATAGAAAAAGC-ATTGATGATTTAAAGAAATTTAAGAATGATTTTAGGTTTACAGCCAAGT 360
QY 2587 TCTACTTCTGAATCTATGATCAACTACTATAAACAATAAGAGGAGGAGAAAAATCAAGA 2646
Db 361 TCTACTTCTGAATCTATGATCAACTACTATAAACAATAAGAGGAGGAGAAAAATCAAGA 420
QY 2647 GATTTGATCAAAAGACAAAAATTGAACCAAGTGTCTAGGATTTCTTATTGAAATAGCAGC 2706
Db 421 GATTTGATCAAAAGACAAAAATTGAACCAAGTGTCTAGGATTTCTTATTGAAATAGCAGC 480
QY 2707 AGCACTGTACAGTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2766
Db 481 AGCACTGTACAGTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 540
QY 2767 AGTAACACCGGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGGTTCAGACTTCCAGC 2826
Db 541 AGTAACACCGGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGGTTCAGACTTCCAGC 600
QY 2827 CCAGCATGTAAACAAGAGAGAACACGATAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGTT 2886
Db 601 CCAGCATGTAAACAAGAGAGAACACGATAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGTT 660
QY 2887 AGGAATCAACATTTGAATCCCAATGCCAATGCCAATGCCAATGCCAATGCCAATGCCAATGCCA 2946
Db 661 AGGAATCAACATTTGAATCCCAATGCCAATGCCAATGCCAATGCCAATGCCAATGCCAATGCCA 720
QY 2947 AGCCCTTCTACTACCCCACTTCCACCTCGGCTCAAGCAGCAACCTAGCCCCATCTATGGTG 3006
Db 721 AGCCCTTCTACTACCCCACTTCCACCTCGGCTCAAGCAGCAACCTAGCCCCATCTATGGTG 780
QY 3007 GGTTCATCAACAGCAGCAACTCCAGTTTATCTACGCTGTTTGTGTTTGCACCAATATGATG 3066
Db 781 GGTTCATCAACAGCAGCAACTCCAGTTTATCTACGCTGTTTGTGTTTGCACCAATATGATG 840
QY 3067 TATCCAGTCCAGTGGAGCCCGAGCGGTGCAACCTTTTATACCCAA-TACCTATGAGCCCAT 3125
Db 841 TATMCAGTCCAGTGGAGCCCGAGCGGTGCAACCTTTTATACCCAA-TACCTATGAGCCCAT 900
QY 3126 GCCAGTGAATCAAGCCCAAGACATATAGAGCAG 3157
Db 901 GCCAGTGAATCAAGCCCAAGACATATAGAGCAG 932

RESULT 6
CN646951
LOCUS
DEFINITION ILLUMIGEN_MCO_27453 Katze_MBR Macaca mulatta cDNA clone IB1UW:8082
```


5' similar to Bases 1 to 877 highly similar to human SCA2
(Hs.76253), mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

EN646951
EN646951.1 GI:47160394

EST.
Macaca mulatta (rhesus monkey)

Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE
AUTHORS

1 (bases 1 to 977)
Magnes, C.L., Fellin, P.C., Thomas, M.J., Korth, M.J., Agy, M.B.,
Prohl, S.C., Fitzgibbon, M., Scherer, C.A., Miner, D.G., Katze, M.G. and
Iadonato, S.P.

TITLE
Analysis of the Macaca mulatta transcriptome and the sequence
divergence between Macaca and human

JOURNAL
PUBMED
COMMENT

Genome Biol. 6 (7), R60 (2005)

15998449

Contact: C. Magnes

Illumigen Biosciences Inc.

2203 Airport Way S, Suite 450, Seattle, WA 98134, USA

Tel: 2063780400

Fax: 2063780408

Email: cmagnes@illumigen.com

Sequenced on 2004.03.18. 774 Q20 bases.

PCR Primers

FORWARD: CCCTCACTAAAGGGGAACAAA

BACKWARD: CACTATAGGCGAATCGGTA

Insert Length: 977 Std Error: 0.00

Plate: CL000183 row: A column: 06

Seq primer: CCCTCACTAAAGGGGAACAAA

POLYA=No.

FEATURES
source

Location/Qualifiers

1..977

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="Indian"

/db_xref="taxon:9544"

/clone="IBIUM:8082"

/sex="female"

/dev_stage="adult"

/lab_host="E. coli SOLR"

/clone_lib="Katze_MMR"

/note="Organ: brain; Vector: Uni-ZAP XR; Site 1: Ecor I;

Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis

kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold

Cloning Kit (Catalog #200450)"

ORIGIN

Query Match 19.0%; Score 850.6; DB 7; Length 977;
Best Local Similarity 98.9%; Pred. No. 3.4e-183;
Matches 867; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 3025 CCAGTTTATCTACGCTGTTGTTTGGACCAAAATATGATGATCCAGTCCAGTGAGC 3084

Db 1 CCAGTTTATCTACGCTGTTGTTTGGACCAAAATATGATGATCCAGTCCAGTGAGC 60

QY 3085 CCAGCGGTGCACCTTTATACCAATACCTATGACGCCCATGCCAGTGAATCAAGCCCAAG 3144

Db 61 CCAGCGGTGCACCTTTGTGACCAATACCTATGACGCCCATGCCAGTGAATCAAGCCCAAG 120

QY 3145 ACATATAGCAGCAGTACCAAAATATGCCCAACAGCGGCAAGACCATCATCAGAGTGCC 3204

Db 121 ACATATAGCAGCAGTACCAAAATATGCCCAACAGCGGCAAGACCATCATCAGAGTGCC 180

QY 3205 ATGATGCAACCGCTCAGCAGCGGGCCCGGATTCAGCCACCCAGCAGCTTACTCC 3264

Db 181 ATGATGCAATCCAGCTCAGCAGCGGGCCCGGATTCAGCCACCCAGCAGCTTACTCC 240

QY 3265 ACGCAATATGTTGCTACAGTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTG 3324

Db 241 ACGCAATATGTTGCTACAGTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTG 300

QY 3325 CCACATTATCAGTCTCAGCATCTCTATGCTCTATAGTCTCTGTAATACAGGGTAATGCTAGA 3384

Db 301 CCACATTATCAGTCTCAGCATCTCTATGCTCTATAGTCTCTGTAATACAGGGTAATGCTAGA 360

QY 3385 ATGATGGCACCACCAACACACAGCCCGCTGGTTAGTATCTTCTTACGCAACTCAGTAC 3444

Db 361 ATGATGGCACCACCAACACACAGCCCGCTGGTTAGTATCTTCTTACGCAACTCAGTAC 420

QY 3445 GGGGCTCATGAGCAGACGATGCGATGATGATGCTGCCAAATTTACCATACAAACAGGAG 3504

Db 421 GGGGCTCATGAGCAGACGATGCGATGATGATGCTGCCAAATTTACCATACAAACAGGAG 480

QY 3505 ACAAGCCCTTCTTCTACTTTGCTTCCAGGGCTCCCTTGTCTACGAGTATGCGCAC 3564

Db 481 ACAAGCCCTTCTTCTACTTTGCTTCCAGGGCTCCCTTGTCTACGAGTATGCGCAC 540

QY 3565 CCTAACGCTACCTGCACCCACATCTCCACACCTCAGCTTACGCTACCCCACTGGA 3624

Db 541 CCTAACGCTACCTGCACCCACATCTCCACACCTCAGCTTACGCTACCCCACTGGA 600

QY 3625 CAGCAGCAAAAGCCAAACATGTTGGAAGTCTCTGTCACCCAGTCTCTTTCAGCACCACATCAG 3684

Db 601 CAGCAGCAAAAGCCAAACATGTTGGAAGTCTCTGTCACCCAGTCTCTTTCAGCACCACATCAG 660

QY 3685 CACCAGCGCCCGCCAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTTCAGCCATTTTACCAC 3744

Db 661 CACCAGCGCCCGCCAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTTCAGCCATTTTACCAC 720

QY 3745 CGGGGCTTGGCCCACTCCACCTCCATGACACCTGCTCCACACCGCAGTTCGCCACAG 3804

Db 721 CGGGGCTTGGCCCACTCCACCTCCATGACACCTGCTCCACACCGCAGTTCGCCACAG 780

QY 3805 AATAGTTTCCCAGCAGCACAACAGTCTCTTTTCAGTCCATCTCTCAGCTTCAGCCG 3864

Db 781 AATAGTTTCCCAGCAGCACAACAGTCTCTTTTCAGTCCATCTCTCAGCTTCAGCCG 840

QY 3865 CGGTATACCAACCCA-CCCCACATGGCCACGTACCT 3900

Db 841 CGCGAAACCAACCCACCCACATGGCCACGTACCT 877

RESULT 7
BM455214 1100 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6405612 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500163
DEFINITION 5', mRNA sequence.
ACCESSION BM455214
VERSION BM455214.1 GI:18504254
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12134 row: k column: 12
High quality sequence stop: 623.

REFERENCE 1 (bases 1 to 1100)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12134 row: k column: 12
High quality sequence stop: 623.

FEATURES
source
1..1100
/organism="Homo sapiens"
/mol_type="mRNA"

```
/db_xref="taxon:9606"
/clone="IMAGE:5500163"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/notes="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 18.4%; Score 842; DB 3; Length 1100;
Best Local Similarity 96.4%; Pred. No. 3.2e-181;
Matches 925; Conservative 0; Mismatches 24; Indels 11; Gaps 6;

QY 380 CTTACAGCTGTTTGTAGCAACGCGGCGGCGGCGGCTTTCGGCCGCGCTCCCGGC 439
DB 1 CTTACAGCTGTTTGTAGCAACGCGGCGGCGGCGGCTTTCGGCCGCGCTCCCGGC 60
QY 440 GCGTCTTGGTCTCGCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCGCTCCCA 499
DB 61 GCGTCTTGGTCTCGCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCGCTCCCA 120
QY 500 GCGCGGCGGCTTCCCGGCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCGCTCC 559
DB 121 GCGCGGCGGCTTCCCGGCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCGCTCC 180
QY 560 GCGCGGCTTCCCGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCG 619
DB 181 GCGCGGCTTCCCGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCG 240
QY 620 AGCGGTGTATGGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCG 679
DB 241 AGCGGTGTATGGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCG 300
QY 680 AGCAA---CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 736
DB 301 AGCAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 360
QY 737 CTGCGGCTTCCCGGCGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCC 796
DB 361 CTGCGGCTTCCCGGCGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCC 420
QY 797 CGCGGCTTCTCGGCTTCTCGGCTTCTCGGCTTCTCGGCTTCTCGGCTTCTCGGCTTCTCG 856
DB 421 CGCGGCTTCTCGGCTTCTCGGCTTCTCGGCTTCTCGGCTTCTCGGCTTCTCGGCTTCTCG 480
QY 857 CGACCTTCCCGGCGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCG 916
DB 481 CGACCTTCCCGGCGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCCCG 540
QY 917 CTGAGTCTAGGATTTCTTTGATGGAATCTATGCAATATGAGATGTTTCTATATCTTA 976
DB 541 CTGAGTCTAGGATTTCTTTGATGGAATCTATGCAATATGAGATGTTTCTATATCTTA 600
QY 977 CATCAGTGTGTGGCTCCAAATGTGAAGTACAAGTGAAGTGAAGTATATCAAGGAG 1036
DB 601 CATCAGTGTGTGGCTCCAAATGTGAAGTACAAGTGAAGTGAAGTATATCAAGGAG 660
QY 1037 TTTTAAACCTTACAGTCCGAGTGTGATTTGGTACTTGTGATCGGCGCATGAGAAAGTA 1096
DB 661 TTTTAAACCTTACAGTCCGAGTGTGATTTGGTACTTGTGATCGGCGCATGAGAAAGTA 720
QY 1097 CAGATCCAGTTCGGGCGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCC 1156
DB 721 CAGATCCAGTTCGGGCGGCGGCTTCCCGGCGGCTTCCCGGCGGCTTTCGGTCTCTTCTCC 780
QY 1157 CAGATTTGTGTGGTACAG-TTTAAAGATATGAGTCCAGTTATGCAAAAGAGATGCT 1215
DB 781 CAGATTTGTGTGGTACAGTTTAAAGATATGAGTCCAGTTATGCAAAAGAGATGCT 840
QY 1216 TTTACTGACTGTCTATCAGTGC-TAAAGTGAATGGC-GAACACAAAGAGAGACCTGG 1273
```

```
|||||
841 TTTACTGACTGTCTATCAGTGCCTAAAGTGAATGCGGACCCCAAGAGAGGACCTGG 900
1274 A--GCCCTGGGATGAGGTGAA---CTCACGCCAATGAGAACTTGAGGCTTTGGAAAA 1328
901 NAGCCCTGGGATGAGGGAACCTCCAGGCCAATTAGGGAACCTTGAAGCTTTTGGNAA 960

BQ716397 859 bp mRNA linear EST 16-JUL-2002
AGENCOURT 8477984 Lupski sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6196657 5', mRNA sequence.
BQ716397
VERSION BQ716397.1 GI:21855294
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 859)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13605 row: h column: 02
High quality sequence stop: 659.

FEATURES
source
1..859 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6196657"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski sympathetic trunk"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGTCCG-3' and
5'-GACTAGTCTAGATCGGAGCGGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

ORIGIN
Query Match 18.4%; Score 822.8; DB 5; Length 859;
Best Local Similarity 98.3%; Pred. No. 7.6e-177;
Matches 843; Conservative 0; Mismatches 12; Indels 3; Gaps 1;

QY 1823 CTTGCGCATCTCTCTCTCGCCACCTTCTCGTACCAAGTCAGGTCCCACTCTTTC 1882
DB 1 CTTGCGCATCTCTCTCTCGCCACCTTCTCGTACCAAGTCAGGTCCCACTCTTTC 60
QY 1883 CACTCGGCGAGCACCCCTACAGCGCGCCCTCGAGGCCCTTCGCGCCATCCAGAC 1942
DB 61 CACTCGGCGAGCACCCCTACAGCGCGCCCTCGAGGCCCTTCGCGCCATCCAGAC 120
QY 1943 CCGGCTCTACCCCTCTGCTCATGTTCTCAGCTCTGCTCTACTATATCCATAACGCA 2002
```

Db 121 CCCCGTCTCACCCCTCTGCTCATGTTCTCCAGTCTCCTGTCTCTACTATGCTTAACGCA 180
Qy 2003 TGTCTTTCAGAGGGCTTCAAGGATGTCCCAAGGCCCCAGGACATCTCTCGAAATCACA 2062
Db 181 TGTCTTTCAGAGGGCTTCAAGGATGTCCCAAGGCCCCAGGACATCTCTCGAAATCACA 240
Qy 2063 GAGTTTCTGCTGGAGGGTTCATATCAGTGGCTAGAAATTTGTATCCCAACACCCAC 2122
Db 241 GAGTTTCTGCTGGAGGGTTCATATCAGTGGCTAGAAATTTGTATCCCAACACCCAC 300
Qy 2123 CCAGTGAAGCAGCTACTCTCCAGTAGCAAGACAGTCCCTCGGGGGAGAGCTGGTTCAT 2182
Db 301 CCAGTGAAGCAGCTACTCTCCAGTAGCAAGACAGTCCCTCGGGGGAGAGCTGGTTCAT 360
Qy 2183 CAGTGGTCAAGTGGGTTCAGATTTATCCCTTAATAACTATAGACCCAGGCTCCCGACAG 2242
Db 361 CAGTGGTCAAGTGGGTTCAGATTTATCCCTTAATAACTATAGACCCAGGCTCCCGACAG 420
Qy 2243 AGAACAGTATTGGAAATACCCCAAGTGGGCCAGTCTTGTCTCTCCCAAGCTGGTATTA 2302
Db 421 AGAACAGTATTGGAAATACCCCAAGTGGGCCAGTCTTGTCTCTCCCAAGCTGGTATTA 480
Qy 2303 TTCCAACTGAAGCTGTGTGCTATTCAGCTGCTATTCAGCTGCTATCTAGGCTGCTAGTCTG 2362
Db 481 TTCCAACTGAAGCTGTGTGCTATTCAGCTGCTATTCAGCTGCTATCTAGGCTGCTAGTCTG 540
Qy 2363 CATCGAACAGAGCTGTGTACCCCTTCTAGTGGGCTAAAGATTTCCAGGCTTCAAGATCAGA 2422
Db 541 CATCGAACAGAGCTGTGTACCCCTTCTAGTGGGCTAAAGATTTCCAGGCTTCAAGATCAGA 600
Qy 2423 GGCAGAACTCTCTCGAGGAAATAAGAAATATTAAACCCCAATGAACATCACCCTAGCT 2482
Db 601 GGCAGAACTCTCTCGAGGAAATAAGAAATATTAAACCCCAATGAACATCACCCTAGCT 660
Qy 2483 TCTCAAAAGCTCAAAACAAAGGTATATCACAGTGTGTTTCTGAACATAGAAACAGATTG 2542
Db 661 TCTCAAAAGCTCAAAACAAAGGTATATCACAGTGTGTTTCTGAACATAGAAACAGATTG 720
Qy 2543 ATGATTTAAAGAAATTTAAGATGATTTTAGGTTACAGCAAGTGTCTACTTCTGAATCTA 2602
Db 721 ATGATTTAAAGAAATTTAAGATGATTTTAGGTTACAGCAAGTGTCTACTTCTGAATCTA 780
Qy 2603 TGGATCAACTACTAAACAAATAATAGAGGGAGAGAAATCAAGAGA----TTTGATCAAG 2659
Db 781 TGGATCAACTACTAAACAAATAATAGAGGGAGAGAAATCCAGGAGATTTGGATCAAG 840
Qy 2660 ACAAAATTGAACCAAGTG 2677
Db 841 ACCAAATTGAACCAAGTG 858

RESULT 9
CA489164
LOCUS
DEFINITION AGNCOURT_10853363 MAPcL Homo sapiens cDNA clone IMAGE:6721472 5',
mRNA sequence.
ACCESSION CA489164
VERSION CA489164.1 GI:24951955
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 889)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-romail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM14282 row: k column: 08
High quality sequence stop: 683.
Location/Qualifiers

FEATURES

source

1. .889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:6721472"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HMEL, LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPcL"
/note="Vector: PCMV-SPORT6; Site_1: EcoRV; Site_2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Eglund, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan. Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN

Query Match 18.3%; Score 819.2; DB 6; Length 889;
Best Local Similarity 98.2%; Pred. No. 5.1e-176;
Matches 851; Conservative 0; Mismatches 9; Indels 7; Gaps 2;
Qy 2373 AGCTGTACCCCTTCTAGTGGCTTAAAGATTCAGGCTTCAAGATCAGAGCGAAGCTC 2432
Db 1 AGCTGTACCCCTTCTAGTGGCTTAAAGATTCAGGCTTCAAGATCAGAGCGAAGCTC 60
Qy 2433 TCCTCGAGGGAAATAAGAAATATTAAACCCCAATGAACATCAGATTGATGATTTAAA 2492
Db 61 TCCTCGAGGGAAATAAGAAATATTAAACCCCAATGAACATCAGATTGATGATTTAAA 120
Qy 2493 TGAACCAAAAGGTATATCACAGTGTGTTTCTGAACATAGAAAACAGATTGATGATTTAAA 2552
Db 121 TGAACCAAAAGGTATATCACAGTGTGTTTCTGAACATAGAAAACAGATTGATGATTTAAA 180
Qy 2553 GAAATTTAAGAAATGATTTTAGGTTACAGCAAGTGTCTACTTCTGAATCTATGGATCAACT 2612
Db 181 GAAATTTAAGAAATGATTTTAGGTTACAGCAAGTGTCTACTTCTGAATCTATGGATCAACT 240
Qy 2613 ACTAAACAAATAATAGAGGGAGAGAAATCAAGAGATTTGATCAAGACAAATTTGAACC 2672
Db 241 ACTAAACAAATAATAGAGGGAGAGAAATCAAGAGATTTGATCAAGACAAATTTGAACC 300
Qy 2673 AAGTCTAAGGATTTCTTTTCATTGAAAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAA 2732
Db 301 AAGTCTAAGGATTTCTTTTCATTGAAAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAA 360
Qy 2733 GCCGAATAGCCCGCAGCATTTCCCTTCTTAATCTTAGTAAACGAGGACAAAGAGGGGACC 2792
Db 361 GCCGAATAGCCCGCAGCATTTCCCTTCTTAATCTTAGTAAACGAGGACAAAGAGGGGACC 420
Qy 2793 TGAGTCTACTTTCCCAAGGGGTTCCAGCTTCCAGCCAGCATGTAAACAAAGAAAGACGA 2852
Db 421 TGAGTCTACTTTCCCAAGGGGTTCCAGCTTCCAGCCAGCATGTAAACAAAGAAAGACGA 480
Qy 2853 TAAGGAAGAGAAAGAACGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAAATCCCAATGC 2912
Db 481 TAAGGAAGAGAAAGAACGCGAGCTGAGCAAGTTAGGAAATCAACATTTGAAATCCCAATGC 540
Qy 2913 AAAGAGTTCAACCCACGTTCTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACC 2972
Db 541 AAAGAGTTCAACCCACGTTCTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACC 600
Qy 2973 TCGGCTCAACGACAACTAGCCCATCTATGGTGGTTCATCAACAGCAACCTCCAGTTTA 3032
Db 601 TCGGCTCAACGACAACTAGCCCATCTATGGTGGTTCATCAACAGCAACCTCCAGTTTA 660

QY	3234	ACGATTGCGACGCCACCCACAGCTTACTCCACGCAATATGTTGCTCAGTCTCCTCAGCA	3293
Db	1	ACCGATTGCGACGCCACCCACAGCTTACTCCACGCAATATGTTGCTCAGTCTCCTCAGCA	60
QY	3294	GTTCCCAATCAGCCCTTGTTCAGCATCTGCCACATTAATCAGTCTCAGCATCTCATGT	3353
Db	61	GTTCCCAATCAGCCCTTGTTCAGCATCTGCCACATTAATCAGTCTCAGCATCTCATGT	120
QY	3354	CTATAGTCTCTGTAATAACAGGTAATCTAGAAATGATGGCACCAACACACGCCCCAGCC	3413
Db	121	CTATAGTCTCTGTAATAACAGGTAATCTAGAAATGATGGCACCAACACACGCCCCAGCC	180
QY	3414	TGTTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATAGACAGCGCATGCGATGTA	3473
Db	181	TGTTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATAGACAGCGCATGCGATGTA	240
QY	3474	TGCATCTCCCAATTAACCATACAAAGAGAGACAAGCCCTTCTTCTACTTTGCCATTTC	3533
Db	241	TGCATCTCCCAATTAACCATACAAAGAGAGACAAGCCCTTCTTCTACTTTGCCATTTC	300
QY	3534	CACGGGCTCCCTTGTCTCAGCAGTATGGCAACCCCTAACGCTTACCTTGCACCCACATACTCC	3593
Db	301	CACGGGCTCCCTTGTCTCAGCAGTATGGCAACCCCTAACGCTTACCTTGCACCCACATACTCC	360
QY	3594	ACACCTCAGCTTCAGTACCCCTTACCTGACAGCAAGCAACATGTTGGAAAGTCA	3653
Db	361	ACACCTCAGCTTCAGTACCCCTTACCTGACAGCAAGCAACATGTTGGAAAGTCA	420
QY	3654	TCTGACCCAGTCTTGTTCAGCAACCATCAGCAACCGCGCCGCGCTCTCCATCTGGC	3713
Db	421	TCTGACCCAGTCTTGTTCAGCAACCATCAGCAACCGCGCGCTCTCCATCTGGC	480
QY	3714	CAGTCCACAGCAGTTCAGCCATTTACACGCGGGCTTGGCGCAACTCCACCTCCAT	3773
Db	481	CAGTCCACAGCAGTTCAGCCATTTACACGCGGGCTTGGCGCAACTCCACCTCCAT	540
QY	3774	GACACCTGCTCCAAACGCGAGTGGCCACAGATAGTTTCCAGCAGCAACAGACTGT	3833
Db	541	GACACCTGCTCCAAACGCGAGTGGCCACAGATAGTTTCCAGCAGCAACAGACTGT	600
QY	3834	CTTTACGATCCATCTCTCAGTTCAGCGGGGTATACCAACCCACCCACATGGCCCA	3893
Db	601	CTTTACGATCCATCTCTCAGTTCAGCGGGGTATACCAACCCACCCACATGGCCCA	659
QY	3894	CGTACCTCAGGCTCATGTACAGTCA-CGAAATGGTTCCTCTCATCCAACTGCCCATGGC	3952
Db	660	CGTACCTCAGGCTCATGTACAGTCA-CGAAATGGTTCCTCTCATCCAACTGCCCATGGC	719
QY	3953	CAATGATGCTAATGACGACAGCAGCACCCGCGGTCCCCAGCGCGCTCGCTCAAAGTG	4012
Db	720	CAATGATGCTAATGACGACAGCAGCACCCGCGGTCCCCAGCGCGCTCGCTCAAAGTG	778
QY	4013	CAGTACGCCATTCAGTCTCGACACAGGGCATTTCCCTATATGACGACCCCTTCAG	4072
Db	779	CAGTACAG-CCATTCAGTCTCGACACAGGGCATTTCCCTATATGACGACCCCTTCAG	836
QY	4073	TACAAGCCCAACCA 4088	
Db	837	TACAAGCCCAACCA 852	

RESULT 13
CD616048/c
LOCUS CD616048 830 bp mRNA linear EST 12-JAN-2004
DEFINITION 5609370H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD616048
VERSION CD616048.1 GI:40264312
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae: Homo.
1 (bases 1 to 830)
Fu.G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
Genomics 84 (1), 205-210 (2004)
15203218
Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
Location/Qualifiers
1..830
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

Query Match	17.1%;	Score 767;	DB 6;	Length 830;
Best Local Similarity	97.9%;	Pred. No. 4.4e-164;		
Matches	798;	Conservative 0;	Mismatches 15;	Indels 2;
				Gaps 2;
QY	3279	CTACAGTCTCTCAGCAGTTCGCCAATCAGCCCTTGTTCAGCATGTGCCACATTATCAGTC	3338	
Db	830	CTACAGTCTCTCAGCAGTTCGCCAATCAGCCCTTGTTCAGCATGTGCCACATTATCAGTC	771	
QY	3339	TCAGCATCTCTCATGTCTATAGTCTCTAATACAGGGTAATGCTAGAAATGATGGCACACC	3398	
Db	770	TCAGCATCTCTCATGTCTATAGTCTCTAATACAGGGTAATGCTAGAAATGATGGCACACC	711	
QY	3399	AACACACGCCCCAG-CCTGGTTTGTAGTATCTTTCAGCAACTCAGTACGGGGCTCATGAGC	3457	
Db	710	AACACACGCCCCAGCCCTGGTTTGTAGTATCTTTCAGCAACTCAGTACGGGGCTCATGAGC	651	
QY	3458	AGACGATCGGATGTATGATGTCTCCCAATTACCATTACAACAGGAGACAGCCCTCTTT	3517	
Db	650	AGACGATCGGATGTATGATGTCTCCCAATTACCATTACAACAGGAGACAGCCCTCTTT	592	
QY	3518	TCTACTTTGCCATTTCCACGGGCTCCCTTGTCTCAGCAGTATGCGCACCTTAAGCTACCC	3577	
Db	591	TCTACTTTGCCATTTCCACGGGCTCCCTTGTCTCAGCAGTATGCGCACCTTAAGCTACCC	532	
QY	3578	TGCACCCACATACTCCACACCCCTCAGCTACCCCTTACCTGGAAGCAGCAAGCC	3637	
Db	531	TGCACCCACATACTCCACACCCCTCAGCTACCCCTTACCTGGAAGCAGCAAGCC	472	
QY	3638	AACATGGTGAAGTATCTCTGACCCAGTCTCTTTCAGCAGCATCAGCAGCCGCC	3697	
Db	471	AACATGGTGAAGTATCTCTGACCCAGTCTCTTTCAGCAGCATCAGCAGCCGCC	412	
QY	3698	AGGCTCTCCATCTGGCCAGTCCACAGCAGTCCAGCATTTTACCACGGGGCTTCGCG	3757	
Db	411	AGGCTCTCCATCTGGCCAGTCCACAGCAGTCCAGCATTTTACCACGGGGCTTCGCG	352	
QY	3758	CAACTCCACCCCTCATGACACCTTCCAAACGCGAGTCCGACAGAAATAGTTTCCAG	3817	
Db	351	CAACTCCACCCCTCATGACACCTTCCAAACGCGAGTCCGACAGAAATAGTTTCCAG	292	
QY	3818	CAGCAACAGACTGTCTTTTACGATCCATCTCTTTCAGCTTTCAGCCGGGTATACCAACC	3877	
Db	291	CAGCAACAGACTGTCTTTTACGATCCATCTCTTTCAGCTTTCAGCCGGGTATACCAACC	232	
QY	3878	CACCCACATGGCCAGTACCTCAGCTCAGTACAGTCCAGGAATGTTCTCTCATC	3937	
Db	231	CACCCACATGGCCAGTACCTCAGCTCAGTACAGTCCAGGAATGTTCTCTCATC	172	
QY	3938	CAACTGCCCATGCGCCAAATGATGCTAATGACGACAGCCACCCGGCGGTCCCGAGCCG	3997	
Db	171	CAACTGCCCATGCGCCAAATGATGCTAATGACGACAGCCACCCGGCGGTCCCGAGCCG	112	

QY 3998 CCCTCGCTCAAAAGTGCACCTACAGCCCATTCAGTCTCGACACAGGCAATTTCCCTTATA 4057
 |||||
 Db 111 CCCTCGCTCAAAAGTGCACCTACAGCCCATTCAGTCTCGACACAGGCAATTTCCCTTATA 52
 |||||
 QY 4058 TGACGCGACCCCTTCAGTACAAAGCCACCAACCAACAG 4092
 |||||
 Db 51 TGACGCGACCCCTTCAGTACAAAGCCACCAACCAACAG 17
 |||||

RESULT 14
 C0886647
 LOCUS
 DEFINITION Boven_14972 normal cattle brain Bos taurus CDNA clone
 RZPDP1056D1525Q 5', mRNA sequence.
 C0886647
 ACCESSION
 VERSION
 C0886647.1 GI:51816932
 KEYWORDS
 EST.
 SOURCE Bos taurus (cow)

ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 Pecora; Bovidae; Bovinae; Bos.

REFERENCE
 AUTHORS Hennig S., Janitz M., Herwig R. and Williams J.
 TITLE Generation, annotation, evolutionary analysis and database
 integration of 14969 cattle EST clusters
 JOURNAL Unpublished (2004)
 COMMENT Contact: Hennig S

laboraty 123, dept. Lehrach
 Max-Planck-Institut fuer Molekulare Genetik
 Innestr.63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1612
 Fax: +49 30 8413 1380
 Email: hennigmolgen.mpg.de

The library was characterised by oligonucleotide fingerprinting
 (ONFP) to reduce sequencing redundancy. According to the ONFP
 procedure, clones that display the same hybridisation matrix with a
 battery of 200 8mer oligonucleotides are grouped into clusters. One
 clone per ONFP cluster was selected for sequencing. cDNA clones and
 filters are distributed via Deutsches Ressourcenzentrum fuer
 Genomforschung GmbH (<http://www.rzpd.de>).

PCR Primers
 FORWARD: 5' CCCAGGCTTACATTTATGCTCCGGCTCG 3' (M13RSP) 5'-seq
 BACKWARD: 5' GCTATTAGCAGCTGGCGAAGGGGATGTG 3' (M13FSP) 3'-seq
 Seq primer: 5'-CCGCTCCGGAATTCGGGT-3' (M13RSP).

FEATURES
 Location/Qualifiers
 1..877
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /clone="RZPDP1056D1525Q"
 /sex="female"
 /tissue_type="brain tissue"
 /dev_stage="adult brain"
 /clone_lib="normal cattle brain"
 /note="Organ: brain; Vector: pSport1; Site 1: NotI;
 Site 2: SalI; Random primed and directionally cloned in
 pSport1 vector using NotI
 5'-PGACTAGTCTTAGATCGCGAGCGCGCC (T)15-3' and SalI 5'-
 TCGACCCAGCGCTCG-3' adapters (Gibco BRL)";

ORIGIN
 Query Match 17.0%; Score 761.2; DB 7; Length 877;
 Best Local Similarity 95.6%; Pred. No. 9.5e-163;
 Matches 805; Conservative 0; Mismatches 33; Indels 4; Gaps 2;

QY 3643 GGTGGAGTATCTTCGACCCAGTCTCTGTTCAGCACCATCAGACAGCGCCCGCAGGCT 3702
 |||||
 Db 34 GGTGGAGGATCTCTCGCCGACGCTCTGTTCAGCACCATCAGACAGCGCCCGCAGGCT 93
 |||||

QY 3703 CTCCTCTCGCAGTCCACAGCAGCAGTCAGCCATTACACGCGGGGCTTCGGCCTACT 3762
 |||||
 Db 94 CTCCTCTCGCAGTCCACAGCAGCAGTCAGCCATTACACGCGGGGCTTCGGCCTACT 153
 |||||

QY 3763 CCACCTCCATGACACACCTGCTCCAAACACGCGAGTCGCCACAGAAATAGTTTCCACGACGA 3822
 |||||
 Db 154 CCACCTCCATGACACACCTGCTCCAAATACACAGTCTCCACAGAAATAGTTTCCACGACGA 213
 |||||
 QY 3823 CAACAGACTGTCTTTAGATCCATCCTTCTCAGCTTCAGCGGGGTATACCAACCCACCC 3882
 |||||
 Db 214 CAGCAGACCGGTCTTTACCAATTCATCCTTCTCATGTTTACGCGGGCTTACCAATCCACCC 273
 |||||
 QY 3883 CACATGGCCACAGTACCTACCTCAGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCAACT 3942
 |||||
 Db 274 CACATGGCCATGATACCTCAGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCAACT 333
 |||||
 QY 3943 GCCATGGCCCAATGATCTAATGACGACACAGCCACCCCGGTGCCAGCGCCCTC 4002
 |||||
 Db 334 GCCATGGCCCAATGATCTAATGACGACACAGCCACCCCGGTGCCAGCGCCCTC 393
 |||||
 QY 4003 GCTCAAGTGCACCTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTCTATATGAG 4062
 |||||
 Db 394 GCTCAAGTGCACCTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTCTATATGAG 453
 |||||
 QY 4063 CACCTTTCAGTACAAAGCCACCAACACAGCAGTTGTAAAGCTGCCCTGGAGGAACCGAA 4122
 |||||
 Db 454 CACCTTTCAGTACAAAGCCACCAACACAGCAGTTGTAAAGCTGCCCTGGAGGAACCGAA 513
 |||||
 QY 4123 AGGCCAAATTCCTCTCCTCTTCTACTGCTTCTACCAACTGGAAGCAAGAACTAGAA 4182
 |||||
 Db 514 AGGCCAAATTCCTCTCCTCTTCTACTGCTTCTACCAACTGGAAGCAAGAACTAGAA 573
 |||||
 QY 4183 TTTCAATTTATTTGTTTTTAAATATATATGTTGTTTGTAAACATCCCAATAGGAATG 4242
 |||||
 Db 574 TTTCA-ATTAATTTGTTTTTGAATAAATATATGTTGTTTGTAAACATCCCAATAGGAATG 632
 |||||
 QY 4243 CTAACAGTTCACCTTGCAAGAGATACCTTGACCGAGTAGAGGCAATTTAGGAACCTTGG 4302
 |||||
 Db 633 CTAACAGTTCACCTTGCAAGAGATATTTGACCGAGTAGAGGCAATTTAGGAACCTTGG 692
 |||||
 QY 4303 GGCTATTCCTAATTTCCATATGCTGTTTCAAGTCCCGCAGGTACCCAGCTCTGCTTGC 4362
 |||||
 Db 693 GGCTGTTCCATAATTCATATGCTGTTTCCAAGTCCCGCAAGTACCCAGCTCTGCTTGC 752
 |||||
 QY 4363 CGAACTCGAAGTATTTATTTTATTAACCTTGAAGTCATG---AACACATCACT 4419
 |||||
 Db 753 TCAAACTCGAAGTATTTATTTATTTATGACCTTTGAAGTATGAAACAACATCACT 812
 |||||
 QY 4420 AGCAAAAGAGTAACAAGAGTATTTCTTGTGCTGCTATTACTGCTAAATAAAAAAAAA 4479
 |||||
 Db 813 AGCAAAAGAGTAACAAGAGTATTTCTTGTCTGCTATTACCGCTAATAAAAAAAAA 872
 |||||
 QY 4480 AA 4481
 ||
 Db 873 AA 874

RESULT 15
 CA775370/c
 LOCUS
 DEFINITION io88f10.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6133411 3',
 similar to TR:Q99700 Q99700 ATAXIN-2. ;, mRNA sequence.
 ACCESSION CA775370
 VERSION CA775370.1 GI:26012904
 KEYWORDS
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 784)
 Melton D., Brown J., Kenty G., Permutt A., Lee C., Kaestner K.,
 Lemishka I., Scarce M., Brestelli J., Gradwohl G., Clifton S.,
 Hillier L., Marra M., Pape D., Wylie T., Martin J., Blistain A.,
 Schmitt A., Theising B., Ritter E., Ronko I., Bennett J.,
 Cardenas M., Gibbons M., McCann R., Cole R., Teagareishvili R.,

Williams, T., Jackson, Y. and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohp.harvard.edu
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Dr. Hiroshi Inoue
 (hinoue@im.wustl.edu)
 Seq primer: -40UP from Gibco
 High quality sequence stop: 465.
 Location/Qualifiers

1. .784
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6133411"
 /tissue_type="Purified pancreatic islet"
 /lab_host="Dhi08"
 /clone_lib="HR85 islet"
 /notes="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
 NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
 Size-selected on agarose gel. Average insert size ~1kb. 5'
 XhoI site was destroyed after directional cloning.
 Amplified once. Contact information: Hiroshi Inoue, MD,
 Metabolism Div. (Alan Permutt Lab), Washington University
 School of Medicine, Box 8127, 660 South Euclid Ave., St.
 Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
 314-362-1916, Fax: 314-747-2692."

FEATURES

source

ORIGIN

Query Match 16.8%; Score 753.2; DB 6; Length 784;
 Best Local Similarity 97.4%; Pred. No. 6.3e-161;
 Matches 764; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

3685 CACAGGCGCGCCAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTTCACCATTTACCAC 3744
 |||||
 784 CACCAGGGCGCCAGGTTCTGCTATTTGGCCAGTCCCAACAGCCAGTAGCCCTTTACCAC 725
 |||||
 3745 GCGGGGCTTCGCGCACTCCACCTCCATGACACCTGCTCCCAACAGCAGTCCGCCACAG 3804
 |||||
 724 GCGGGGCTTCGCGCACTCCACCTCCATGACACCTGCTCCCAACAGCAGTCCGCCACAG 665
 |||||
 3805 AATAGTTTCCAGCAGCACAACAGACTGTCTTTACGATCCATCTCTTCACGTTTCAGCCG 3864
 |||||
 664 AATAGTTTCCAGCAGCACAACAGACTGTCTTTAGCATCCATCTCTTCAGTTTCAGCCG 605
 |||||
 3865 GCGTATACCAACCCACCCCAATGGCCCACTGCTACCTCAGGCTCATGTACAGTAGGAATG 3924
 |||||
 604 GCGTATACCAACCCACCCCAATGGCCCACTGCTACCTCAGGCTCATGTACAGTAGGAATG 545
 |||||
 3925 GTTCCTTCTCATCACTGCCCACTGCCCAATGATGCTAATGACACACAGCCACCCGGC 3984
 |||||
 544 GTTCCTTCTCATCACTGCCCACTGCCCAATGATGATGATGATGATGATGATGATGATGATG 485
 |||||
 3985 GGTCCCGAGGCGCGCTCGCTCAAAGTGCACTACAGCCCATTTCCAGTCTCGACAACAGCG 4044
 |||||
 484 GGTCCCGAGGCGCGCTCGCTCAAAGTGCACTACAGCCCATTTCCAGTCTCGACAACAGCG 425
 |||||
 4045 CATTTCCCTTATAGACGACCCCTTCAGTACAGCCCAACCAACAGCAGTTGTAAAGGC 4104
 |||||
 424 CATTTCCCTTATATGACGACCCCTTCAGTACAGCCCAACCAACAGCAGTTGTAAAGGC 365
 |||||
 4105 TGCCCTCGAGGAACCGAAAGCCAAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4164
 |||||
 364 TGCCCTCGAGGAACCGAAAGCCAAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 305

QY 4165 AAGCAGAGAACTAGAAATTTTCAATTTATTTTGTATTTTAAATATATATATGTTGTTCTTG 4224
 |||||
 DB 304 AAGCAGAGAACTAGAAATTTTCAATTTATTTTGTATTTTAAATATATATATGTTGTTCTTG 2445
 |||||
 QY 4225 TAAACATCCCAATAGGAATGCTAAACAGTTCACATTGCAAGTGAAGATACCTTGGACCGAGTAGA 4284
 |||||
 DB 244 TAAACATCCCAATAGGAATGCTAAACAGTTCACATTGCAAGTGAAGATACCTTGGACCGAGTAGA 185
 |||||
 QY 4285 GGCATTTAGGAACCTTGGGGGCTATTCCATAATTCATATGCTGTTTCAGAGTCCCGCAGG 4344
 |||||
 DB 184 GGCATTTAGGAACCTTGGGGGCTATTCCATAATTCATATGCTGTTTCAGAGTCCCGCAGG 125
 |||||
 QY 4345 TACCCAGCTCTGCTTGGCGAACTGGAAGTATTATTTTAAATACCCCTTGAAGTC 4404
 |||||
 DB 124 TACCCAGCTCTGCTTGGCGAACTGGAAGTATTATTTTAAATACCCCTTGAAGTC 65
 |||||
 QY 4405 ATGAACACATCAGCTAGCAAAAGAAAGTAAACAGAGTATTCTTGTGCTTATTACTGCTAA 4464
 |||||
 DB 64 ATGAACACATCAGCTAGCAAAAGAAAGTAAACAGAGTATTAAAGGTGCTATTACTGCTAA 5
 |||||
 QY 4465 AAAAA 4468
 |||||
 DB 4 AAAAA 1

Search completed: December 23, 2005, 05:02:06
 Job time : 15487 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 17:59:31 ; Search time 776 Seconds
(without alignments)
10264.502 Million cell updates/sec

Title: US-10-802-228-1
Perfect score: 4481
Sequence: 1 acccccggagaaagcaaccacccca.....taaaaaaaaaaaaaaaaaaaaaa 4481

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/1_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/H_COMB.seq.*
6: /cgn2_6/prodata/1/ina/PTUS_COMB.seq.*
7: /cgn2_6/prodata/1/ina/PP_COMB.seq.*
8: /cgn2_6/prodata/1/ina/RE_COMB.seq.*
9: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4481	100.0	4481	3	US-09-041-886-18
2	4481	100.0	4481	3	US-09-648-281-1
3	4481	100.0	4481	3	US-09-707-919A-20
4	4481	100.0	4481	3	US-09-083-268-2
5	4481	100.0	4481	3	US-08-981-998A-2
6	4447	99.2	4484	3	US-09-949-016-4498
7	2957	66.0	3798	3	US-08-981-998A-4
8	935.2	20.9	1257	3	US-09-648-281-11
9	935.2	20.9	1257	3	US-09-083-268-4
10	935.2	20.9	1257	3	US-08-981-998A-19
11	870	19.4	151088	3	US-09-949-016-16240
12	571.2	12.7	623	3	US-09-043-303-5
13	497.4	11.1	516	3	US-09-083-268-1
14	497.4	11.1	516	3	US-08-981-998A-1
15	455.8	10.2	459	3	US-09-707-919A-13
16	412.4	9.2	49677	3	US-09-949-002-817
17	354.2	7.9	355	3	US-09-043-303-1
18	313.2	7.0	47677	3	US-09-949-002-668
19	234	5.2	601	3	US-09-949-016-160682
20	183	4.1	183	3	US-09-707-919A-16
21	177	4.0	177	3	US-09-707-919A-17
22	169	3.8	169	3	US-09-707-919A-15
23	167	3.7	205	3	US-09-043-303-3
24	149	3.3	601	3	US-09-949-016-160680

C	25	149	3.3	601	3	US-09-949-016-160681	Sequence 160681
	26	106	2.4	106	3	US-09-707-919A-14	Sequence 14, Appl
	27	99.4	2.2	1614	3	US-09-616-289-45	Sequence 45, Appl
	28	99.4	2.2	1614	3	US-09-976-740-45	Sequence 45, Appl
	29	99.4	2.2	12425	3	US-09-616-289-50	Sequence 50, Appl
	30	99.4	2.2	12425	3	US-09-976-740-50	Sequence 50, Appl
	31	98.8	2.2	2561	3	US-09-616-289-48	Sequence 48, Appl
	32	98.8	2.2	2561	3	US-09-976-740-48	Sequence 48, Appl
	33	98	2.2	98	3	US-09-707-919A-18	Sequence 18, Appl
	34	94.4	2.1	15695	3	US-09-949-016-15644	Sequence 15644, A
	35	93	2.1	3169	3	US-09-620-312D-713	Sequence 713, Appl
	36	87.8	2.0	477	3	US-09-135-994-1	Sequence 1, Appl
	37	87.8	2.0	477	3	US-09-684-843A-1	Sequence 1, Appl
	38	87	1.9	601	3	US-09-949-016-160684	Sequence 160684
C	39	86.2	1.9	8438	2	US-07-945-283-1	Sequence 1, Appl
	40	86.2	1.9	37875	3	US-09-949-016-13182	Sequence 13182, A
	41	85	1.9	12001	2	US-08-458-568A-11	Sequence 11, Appl
	42	81.4	1.8	10348	2	US-08-457-273B-41	Sequence 41, Appl
	43	81.4	1.8	10348	3	US-08-556-419-13	Sequence 13, Appl
	44	81.4	1.8	10348	3	US-09-041-886-14	Sequence 14, Appl
	45	81.4	1.8	10366	2	US-08-246-982A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-041-886-18
; Sequence 18, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4481 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 163..4099
US-09-041-886-18
Query Match 100.0%; Score 4481; DB 3; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ACCCCGAGAAAGCAACCCAGCGCCGCGCCGCTCTCACGTGTCCTCTCCGGGCCCGGG	60						
Db	1	ACCCCGAGAAAGCAACCCAGCGCGCGCCGCTCTCACGTGTCCTCTCCGGGCCCGGG	60						
QY	61	GCCACCTCACGTTCTGCTTCGGTCTGACCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG	120						
Db	61	GCCACCTCACGTTCTGCTTCGGTCTGACCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG	120						
QY	121	CACCTCCGCTCCACCGCGCGCTCCGCGCGCGCCCTCCGATCGATCGCTCAGCGCGCGCA	180						
Db	121	CACCTCCGCTCCACCGCGCGCTCCGCGCGCGCCCTCCGATCGCTCAGCGCGCGCA	180						
QY	181	GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCCCTTCGCGCGACCGCAGGTGGGCC	240						
Db	181	GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCCCTTCGCGCGACCGCAGGTGGGCC	240						
QY	241	GGGTGGCGCTCGCTCCAGGGCGCGCGCGCGGAGCGGGCGGGCGGTGGCGCGCC	300						
Db	241	GGGTGGCGCTCGCTCCAGGGCGCGCGCGCGGAGCGGGCGGGCGGTGGCGCGCC	300						
QY	301	CGGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCCCTCCCTCCCGG	360						
Db	301	CGGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCGCGCCCTCCCTCCCGG	360						
QY	361	CAGAGCTCGCCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGCGCGCGCGCGCG	420						
Db	361	CAGAGCTCGCCTCCCTCCGCGCTCAGACTGTTTGGTAGCAACGGCGCGCGCGCGCG	420						
QY	421	TTTCCGCGCGGCTCCGCGGGCTCCTTGGTCTCGCGGGCGCTCCCGCCCTTCGTGTC	480						
Db	421	TTTCCGCGCGGCTCCGCGGGCTCCTTGGTCTCGCGGGCGCTCCCGCCCTTCGTGTC	480						
QY	481	GTCTTCTCCCTCCGCGCGCGCGCGCCCTCCGCGCGCGCGCGCGCGCCCTCCCGG	540						
Db	481	GTCTTCTCCCTCCGCGCGCGCGCGCCCTCCGCGCGCGCGCGCGCGCCCTCCCGG	540						
QY	541	CTCGCGCGCGGTGCGTCCCGCGCGGTTCGCGGTCCTTGGCGCGCGCGCTCCCGGC	600						
Db	541	CTCGCGCGCGGTGCGTCCCGCGCGGTTCGCGGTCCTTGGCGCGCGCGCTCCCGGC	600						
QY	601	TGTCCCGCGCGGTGCGCGCGGTATGGGCGCTTACCATGCTCGTAGGCGCCCGAG	660						
Db	601	TGTCCCGCGCGGTGCGCGCGGTATGGGCGCTTACCATGCTCGTAGGCGCCCGAG	660						
QY	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG	720						
Db	661	CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG	720						
QY	721	CAGCGCGCGCGGTGCGGCTTCCGCAAGCCGCGCGCGCGCGCTTCTAGCGTGG	780						
Db	721	CAGCGCGCGCGGTGCGGCTTCCGCAAGCCGCGCGCGCGCGCTTCTAGCGTGG	780						
QY	781	CCGCGCGCGCGCTTCCGCGCTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT	840						
Db	781	CCGCGCGCGCGCTTCCGCGCTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT	840						
QY	841	TCCTCGGTGGTCCGCGCGACCTCCGCGCGCGCGCGCGCGCGCTGGGCGAGGTCGAAC	900						
Db	841	TCCTCGGTGGTCCGCGCGACCTCCGCGCGCGCGCGCGCGCGCTGGGCGAGGTCGAAC	900						
QY	901	AGTAACAAAGGACTCGCTCAGTCTACGATTTCTTTTGGTGAATCTATGCAAAATATGAG	960						
Db	901	AGTAACAAAGGACTCGCTCAGTCTACGATTTCTTTTGGTGAATCTATGCAAAATATGAG	960						
QY	961	ATGGTTCAATATCTTACATCAGTTGTCGCTCCAAATGTGAAGTACAAGTGAATAATGGA	1020						
Db	961	ATGGTTCAATATCTTACATCAGTTGTCGCTCCAAATGTGAAGTACAAGTGAATAATGGA	1020						
QY	1021	GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAGTGTGATTTGCTACTTGATGCC	1080						
Db	1021	GGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAGTGTGATTTGCTACTTGATGCC	1080						

QY	1081	GCAATGAGAAAGGTACAGAAATCCAGTTCGGGGCGAAACGTGAAGAAATATGAGAGT	1140
Db	1081	GCAATGAGAAAGGTACAGAAATCCAGTTCGGGGCGAAACGTGAAGAAATATGAGAGT	1140
QY	1141	ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTACAGTTTAAAGATATGGACTCCAGTTAT	1200
Db	1141	ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTACAGTTTAAAGATATGGACTCCAGTTAT	1200
QY	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCGAACACAAA	1260
Db	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGGCGAACACAAA	1260
QY	1261	GAGAAGGACCTGGAGCCCTGGGATGCGAGTGAATTCAGAGCAATAGAGAACTTTGAGCT	1320
Db	1261	GAGAAGGACCTGGAGCCCTGGGATGCGAGTGAATTCAGAGCAATAGAGAACTTTGAGCT	1320
QY	1321	TTGAAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA	1380
Db	1321	TTGAAAAATGACGTATCTAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA	1380
QY	1381	AATTATGTTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTACAA	1440
Db	1381	AATTATGTTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTACAA	1440
QY	1441	AGAGATAACTCAGAAAGAAATTTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAA	1500
Db	1441	AGAGATAACTCAGAAAGAAATTTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAA	1500
QY	1501	ATTGAGTCAAGTGGCGAGTACAAAGCTCGAGTGGCCCTGGAAAAATGATAGAGTGGAG	1560
Db	1501	ATTGAGTCAAGTGGCGAGTACAAAGCTCGAGTGGCCCTGGAAAAATGATAGAGTGGAG	1560
QY	1561	GAAGAAAAATACAGCAGAGTTTCAGAGAAATTCAGAGTGAACGTGAGGGCACAGATAAAC	1620
Db	1561	GAAGAAAAATACAGCAGAGTTTCAGAGAAATTCAGAGTGAACGTGAGGGCACAGATAAAC	1620
QY	1621	ACTAGGGAAAAATAATATATTTCTCTCTGGACAAAGAAATAGAGAAATCATATCTCTGGGA	1680
Db	1621	ACTAGGGAAAAATAATATATTTCTCTCTGGACAAAGAAATAGAGAAATCATATCTCTGGGA	1680
QY	1681	AGTGGGAGACAGAAATTCACCGGTATGGGCCAGCTGGATCGGGCTCCATGCCATCAAGA	1740
Db	1681	AGTGGGAGACAGAAATTCACCGGTATGGGCCAGCTGGATCGGGCTCCATGCCATCAAGA	1740
QY	1741	TCCACTTCTACACTTCAGATTTTCAACCCGAAATTCGGTTCAGACCAAGAGTGTAGTAAAT	1800
Db	1741	TCCACTTCTACACTTCAGATTTTCAACCCGAAATTCGGTTCAGACCAAGAGTGTAGTAAAT	1800
QY	1801	GGAGTGTTCCTCGCCATCGCCTTGGCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1860
Db	1801	GGAGTGTTCCTCGCCATCGCCTTGGCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1860
QY	1861	CAGTCAAGTTCCT	1920
Db	1861	CAGTCAAGTTCCT	1920
QY	1921	CCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1980
Db	1921	CCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1980
QY	1981	GTCTCTACTATGCTTAAACGATGCTTTTCAAGAGGGCTTCAAGAGTGTCTCCCAAAGGCC	2040
Db	1981	GTCTCTACTATGCTTAAACGATGCTTTTCAAGAGGGCTTCAAGAGTGTCTCCCAAAGGCC	2040
QY	2041	CAGGCACATCTCGAAATCAGAGTTTCTCTCGGAGGGGTTCCATATCCAGTGGCCCTA	2100
Db	2041	CAGGCACATCTCGAAATCAGAGTTTCTCTCGGAGGGGTTCCATATCCAGTGGCCCTA	2100
QY	2101	GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCTCAGTAGCAAGGACCACT	2160
Db	2101	GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCTCAGTAGCAAGGACCACT	2160

QY 2161 CCCTGGGGGAAACGTGGTCATCAGTGGTCAAGTGGGTTCCAAAGATTATCCCTAAAAC 2220
DB 2161 CCCTGGGGGAAACGTGGTCATCAGTGGTCAAGTGGGTTCCAAAGATTATCCCTAAAAC 2220
QY 2221 CATAGACCAGAGTCTCCACAGACAGAACTGTTGGAATACCCAGTGGGCGAGTCTT 2280
DB 2221 CATAGACCAGAGTCTCCACAGACAGAACTGTTGGAATACCCAGTGGGCGAGTCTT 2280
QY 2281 GCTTCTCCCAAGCTGGTATTTCCAACTGAAGCTGTTGCCATGCTATTCACAGTCGA 2340
DB 2281 GCTTCTCCCAAGCTGGTATTTCCAACTGAAGCTGTTGCCATGCTATTCACAGTCGA 2340
QY 2341 TCTCTACGCCCTAGTCTGCTGCA TCGAACAGAGCTGTTACCCCTCTCTAGTGAAGCTAAA 2400
DB 2341 TCTCTACGCCCTAGTCTGCTGCA TCGAACAGAGCTGTTACCCCTCTCTAGTGAAGCTAAA 2400
QY 2401 GATTCAGAGCTTCAAGATCAGAGGAGAACTCTCTGAGGGAATAAAGAAATATTTAA 2460
DB 2401 GATTCAGAGCTTCAAGATCAGAGGAGAACTCTCTGAGGGAATAAAGAAATATTTAA 2460
QY 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAAAA CAAAGGTATATCACAGTGT 2520
DB 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAAAA CAAAGGTATATCACAGTGT 2520
QY 2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGNATGATTTTAGTTACAG 2580
DB 2521 TCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGNATGATTTTAGTTACAG 2580
QY 2581 CCAAGTCTCTACTTCTCAATCTATGATCAACTACTAAACAAATAAGAGGAGGAGAAAA 2640
DB 2581 CCAAGTCTCTACTTCTCAATCTATGATCAACTACTAAACAAATAAGAGGAGGAGAAAA 2640
QY 2641 TCAAGAGATTTGATCAAGACAAAATTGAACCAAGTCTCAAGGATCTTTTCATTGAAAT 2700
DB 2641 TCAAGAGATTTGATCAAGACAAAATTGAACCAAGTCTCAAGGATCTTTTCATTGAAAT 2700
QY 2701 AGCAGAGCAACTGTACAGTGGCAGGAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA 2760
DB 2701 AGCAGAGCAACTGTACAGTGGCAGGAGCAAGCCGAATAGCCCCAGCATTTCCCTTCA 2760
QY 2761 ATACTTGTAGTAAACGAGACAGAGAGGAGCTGAGTCACTTCCCAAGGGTTCAGACT 2820
DB 2761 ATACTTGTAGTAAACGAGACAGAGAGGAGCTGAGTCACTTCCCAAGGGTTCAGACT 2820
QY 2821 TCCAGGCCAGCATGTAAA CAAAGAGAAAGACGATAAGGAGAGAGAAAGACGAGCTGAG 2880
DB 2821 TCCAGGCCAGCATGTAAA CAAAGAGAAAGACGATAAGGAGAGAGAAAGACGAGCTGAG 2880
QY 2881 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCGATCTCTCT 2940
DB 2881 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCGATCTCTCT 2940
QY 2941 CAGCCAAAGCTTCTACTACCCCAACTTCACTCCGCTCAAGCACAACCTAGCCCATCT 3000
DB 2941 CAGCCAAAGCTTCTACTACCCCAACTTCACTCCGCTCAAGCACAACCTAGCCCATCT 3000
QY 3001 ATGGTGGTTCATCAACAGCAACTCCAGTTTATCTCAGCTGTTGTTTGTGACCAAAAT 3060
DB 3001 ATGGTGGTTCATCAACAGCAACTCCAGTTTATCTCAGCTGTTGTTTGTGACCAAAAT 3060
QY 3061 ATGATGTATCCAGTCCAGTGAGCCGAGGCTGGAACCTTTATACCAATACCTATGAGC 3120
DB 3061 ATGATGTATCCAGTCCAGTGAGCCGAGGCTGGAACCTTTATACCAATACCTATGAGC 3120
QY 3121 CCCATGCGAGTGAATCAAGCCAGACATATAGAGCAGTACCAATATGCCCAACAGCGG 3180
DB 3121 CCCATGCGAGTGAATCAAGCCAGACATATAGAGCAGTACCAATATGCCCAACAGCGG 3180
QY 3181 CAAGACAGCATCATCAGAGTGCATGATGACCCAGCGTCAGCAGGGGCCCAACCGATT 3240
DB 3181 CAAGACAGCATCATCAGAGTGCATGATGACCCAGCGTCAGCAGGGGCCCAACCGATT 3240
QY 3241 GCAGGCCACCCACGAGTTTACTCCACGCAATATGTTGCCCTACAGTCTCAGCAGTTCCCA 3300

DB 3241 GCAGGCCACCCACGAGTTTACTCCACGCAATATGTTGCCCTACAGTCTCAGCAGTTCCCA 3300
QY 3301 AATCAGCCCCCTTGTTCAGCATGTGCCAATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
DB 3301 AATCAGCCCCCTTGTTCAGCATGTGCCAATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
QY 3361 CCGTAAATACAGGGTAAATGCTAGATGAGCACCACCAACAGAGCCAGCCTGGTTTA 3420
DB 3361 CCGTAAATACAGGGTAAATGCTAGATGAGCACCACCAACAGAGCCTGGTTTA 3420
QY 3421 GTATCTCTTTCAGCAACTCAGTACGGGCTCATGAGCAGACGCGATGATGATGATG 3480
DB 3421 GTATCTCTTTCAGCAACTCAGTACGGGCTCATGAGCAGACGCGATGATGATGATG 3480
QY 3481 CCCAAATTAACATCAACAAAGGAGACAAGCCCTTCTTTCTACTTTGCCATTTCCACGGC 3540
DB 3481 CCCAAATTAACATCAACAAAGGAGACAAGCCCTTCTTTCTACTTTGCCATTTCCACGGC 3540
QY 3541 TCCCTTGTTCAGCATGTGCCACCTTAACGCTACCTTCACCCACATACTCCACACCT 3600
DB 3541 TCCCTTGTTCAGCATGTGCCACCTTAACGCTACCTTCACCCACATACTCCACACCT 3600
QY 3601 CAGCCTTTCAGCTTACCCCACTGGACAGAGCAAAAGCAACATGGTGAAGTCTATCTGCA 3660
DB 3601 CAGCCTTTCAGCTTACCCCACTGGACAGAGCAAAAGCAACATGGTGAAGTCTATCTGCA 3660
QY 3661 CCCAGTCTCTGTTTCAGCACCATCAGACAGGCGGCCAGGCTCTCCATCTGCCAGTCCA 3720
DB 3661 CCCAGTCTCTGTTTCAGCACCATCAGACACAGGCGGCCAGGCTCTCCATCTGCCAGTCCA 3720
QY 3721 CAGCAGCAGTACGACATTTACACGGGGGCTTGGCGCAACTCCACCTTCATGACACCT 3780
DB 3721 CAGCAGCAGTACGACATTTACACGGGGGCTTGGCGCAACTCCACCTTCATGACACCT 3780
QY 3781 GCCTCCAAACAGCAGTGGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTACG 3840
DB 3781 GCCTCCAAACAGCAGTGGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTACG 3840
QY 3841 ATCCATCTCTCTCAGTTCAGCGGGGTATACAAACCCACCCACATGCCCCACGTACCT 3900
DB 3841 ATCCATCTCTCTCAGTTCAGCGGGGTATACAAACCCACCCACATGCCCCACGTACCT 3900
QY 3901 CAGGCTCATGTACAGTTCAGGAATGTTCTCTCATCAACTGCCCATGGGCAATGATG 3960
DB 3901 CAGGCTCATGTACAGTTCAGGAATGTTCTCTCATCAACTGCCCATGGGCAATGATG 3960
QY 3961 CTAATGACACACAGCACCCCGGCTCCAGGGCGCCCTCGCTCAAAAGTGCATACAG 4020
DB 3961 CTAATGACACACAGCACCCCGGCTCCAGGGCGCCCTCGCTCAAAAGTGCATACAG 4020
QY 4021 CCCATTTCCAGTCTCGAACAACAGCGCATTTCCCTCATATGACGACCCCTTCAGTACAAGCC 4080
DB 4021 CCCATTTCCAGTCTCGAACAACAGCGCATTTCCCTCATATGACGACCCCTTCAGTACAAGCC 4080
QY 4081 CACCACAAACAGCAGTTGTAAGCTCCCTGAGGAAACCGAAAGGCCAAATTCCTCTCTC 4140
DB 4081 CACCACAAACAGCAGTTGTAAGCTCCCTGAGGAAACCGAAAGGCCAAATTCCTCTCTC 4140
QY 4141 CCTTCTACTGCTTCTACCAACTGGAAGACAGAAAACTAGAAATTTCAATTTATTTGTTTT 4200
DB 4141 CCTTCTACTGCTTCTACCAACTGGAAGACAGAAAACTAGAAATTTCAATTTATTTGTTTT 4200
QY 4201 TAAATATATATGTTGATTTCTTTGAACATCCAAATAGGAATGCTAACAGTTCATTTGCAG 4260
DB 4201 TAAATATATATGTTGATTTCTTTGAACATCCAAATAGGAATGCTAACAGTTCATTTGCAG 4260
QY 4261 TGAAGATATCTTGGACCGAGTAGAGCATTTAGGAACTTTGGGGCTTATTTCCATTAATTTCCA 4320
DB 4261 TGAAGATATCTTGGACCGAGTAGAGCATTTAGGAACTTTGGGGCTTATTTCCATTAATTTCCA 4320
QY 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGGCGAAACTGGAAGTATTTT 4380

Db 4321 TATGCTGTTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGC CGAAACTGGAAGTTATTT 4380
Qy 4381 AFTTTTAAATACCCCTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAAGAGT 4440
Db 4381 AFTTTTAAATACCCCTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAAACAAGAGT 4440
Qy 4441 GATTCTGCTGCTATCTACTGCTTAAAAAATAAAAAAAAAAAAAA 4481
Db 4441 GATTCTGCTGCTATCTACTGCTTAAAAAATAAAAAAAAAAAAAA 4481

RESULT 2
US-09-648-281-1
; Sequence 1, Application US/09648281
; Patent No. 6515197
; GENERAL INFORMATION:
; APPLICANT: Pulet, Stefan M.
; TITLE OF INVENTION: Transgenic Animal Model of
; TITLE OF INVENTION: Neurodegenerative Disease and Methods of Use
; FILE REFERENCE: P-CE 4336
; CURRENT APPLICATION NUMBER: US/09/648,281
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163)...(4099)
US-09-648-281-1

Query Match 100.0%; Score 4481; DB 3; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCCCCGAGAAAGCAACCCAGCGCCCGCTCTCTCAAGTGTCCCTCCCGGCCCGGG 60
Db 1 ACCCCCGAGAAAGCAACCCAGCGCCCGCTCTCTCAAGTGTCCCTCCCGGCCCGGG 60

Qy 61 GCCACTCAGCTTCGCTTCGCTCGCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120
Db 61 GCCACTCAGCTTCGCTTCGCTCGCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120

Qy 121 CACCTCCGCTCCACCCCGCGCTCGCGCGCCGCCCTCCGATCGCTCAGCGGCCGCA 180
Db 121 CACCTCCGCTCCACCCCGCGCTCGCGCGCCGCCCTCCGATCGCTCAGCGGCCGCA 180

Qy 181 GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCCGCTTGGCCGACGAGTGGGCC 240
Db 181 GCTCTCGAGTCCCGCGTGGCCACCGAGTCTCGCCGCTTGGCCGACGAGTGGGCC 240

Qy 241 GGGTGGCGTCTGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db 241 GGGTGGCGTCTGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300

Qy 301 CCGGACCGTATCCCTCCCGCGCCCTCCCGCGCCCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 CCGGACCGTATCCCTCCCGCGCCCTCCCGCGCCCGCGCGCGCGCGCGCGCGCGCG 360

Qy 361 CAGAGTCCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGCGCGCG 420
Db 361 CAGAGTCCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAACGGCAACGGCGCGCGCG 420

Qy 421 TTTGGCGCGGCTCCCGCGCGCTTGGTCTCGCGCGCGCTCCCGCGCGCTTGGTCTGTC 480
Db 421 TTTGGCGCGGCTCCCGCGCGCTTGGTCTCGCGCGCGCTCCCGCGCGCTTGGTCTGTC 480

Qy 481 GTCCCTTCTCCCGCTCGCACGCCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCG 540
Db 481 GTCCCTTCTCCCGCTCGCACGCCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCG 540

Qy 541 CTCGGCGCCGCTGCTGCTCCCGCCCGCTTCGCGCGTCTCCTTTGGCGCGCCCGGCTCCCGGC 600
Db 541 CTCGGCGCCGCTGCTGCTCCCGCCCGCTTCGCGCGTCTCCTTTGGCGCGCCCGGCTCCCGGC 600

Qy 601 TGTCTCCCGCCGCGCTGCTATGGGCGCCCTCACCATGTCGCTGAAGCCCGAG 660
Db 601 TGTCTCCCGCCGCGCTGCTATGGGCGCCCTCACCATGTCGCTGAAGCCCGAG 660

Qy 661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Db 661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720

Qy 721 CAGCCGCGCCCGCGCTGCGCAATGTCGCAAGCCCGCGCGCGCGCGCTTCTAGCGTTCG 780
Db 721 CAGCCGCGCCCGCGCTGCGCAATGTCGCAAGCCCGCGCGCGCGCTTCTAGCGTTCG 780

Qy 781 CCGCGCGCGCGCGCTTCCCGCTCTCGCTCTCGCTCTCGCTCTCGCTCTCGCTCTCGCT 840
Db 781 CCGCGCGCGCGCGCTTCCCGCTCTCGCTCTCGCTCTCGCTCTCGCTCTCGCTCTCGCT 840

Qy 841 TCCTCGGTGCTGCGCGGACCTCCGCGCGCGGAGGCGCGCGCTGGGCGAGAGTTCGAAAC 900
Db 841 TCCTCGGTGCTGCGCGGACCTCCGCGCGCGGAGGCGCGCGCTGGGCGAGAGTTCGAAAC 900

Qy 901 AGTAAACAAAGGACTGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAAAATATGAGG 960
Db 901 AGTAAACAAAGGACTGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAAAATATGAGG 960

Qy 961 ATGGTTTCATATACATACATCAGTTGTTGGCTCCAAATGTAAGTACAGTCAAAATATGGA 1020
Db 961 ATGGTTTCATATACATACATCAGTTGTTGGCTCCAAATGTAAGTACAGTCAAAATATGGA 1020

Qy 1021 GGTATATATGAAGGAGTTTTTAAACTTACAGTCCGAAAGTGTGATTTGGTACTTGTATGCC 1080
Db 1021 GGTATATATGAAGGAGTTTTTAAACTTACAGTCCGAAAGTGTGATTTGGTACTTGTATGCC 1080

Qy 1081 GCA CATGAGAAAGATACAGAAATCAGTTTCGGGCGGAAACGTGAAGAAATATGAGAGT 1140
Db 1081 GCA CATGAGAAAGATACAGAAATCAGTTTCGGGCGGAAACGTGAAGAAATATGAGAGT 1140

Qy 1141 ATTTTGTTCATATGCTCAGACTTTGTTGTGTACAGTTTAAAGATATGAGTCCAGTTAT 1200
Db 1141 ATTTTGTTCATATGCTCAGACTTTGTTGTGTACAGTTTAAAGATATGAGTCCAGTTAT 1200

Qy 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGCGGAAACACAAA 1260
Db 1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAAGTGAATGCGGAAACACAAA 1260

Qy 1261 GAGAAAGACCTGGAGCCCTGGGATGCACTCAGCAGCAATGAGGAACTTTGAGGCT 1320
Db 1261 GAGAAAGACCTGGAGCCCTGGGATGCACTCAGCAGCAATGAGGAACTTTGAGGCT 1320

Qy 1321 TTGGAAATGACGTATCTTAATGGATGGATCCCAATGATATGTTTCCGATATAAGTAAAGAA 1380
Db 1321 TTGGAAATGACGTATCTTAATGGATGGATCCCAATGATATGTTTCCGATATAAGTAAAGAA 1380

Qy 1381 AATTATGCTGATGCTTACGATATGATGAGTATTTCTTCTGATACAGTCCCTTAAAG 1440
Db 1381 AATTATGCTGATGCTTACGATATGATGAGTATTTCTTCTGATACAGTCCCTTAAAG 1440

Qy 1441 AGAGATAACTCAGAAAGATTTTAAACGGGAGGAGGCGGAAACAGTTCAGTAAAGAA 1500
Db 1441 AGAGATAACTCAGAAAGATTTTAAACGGGAGGAGGCGGAAACAGTTCAGTAAAGAA 1500

Qy 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATGATGATGAGGTCAG 1560
Db 1501 ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAAATGATGATGAGGTCAG 1560

Qy 1561 GAAGAAAAATACACAGCAGTTCAGAGAAATTCAGTGAACGTGAGGGGCGACAGCATAAAC 1620
Db 1561 GAAGAAAAATACACAGCAGTTCAGAGAAATTCAGTGAACGTGAGGGGCGACAGCATAAAC 1620

Qy 1621 ACTAGGGGAAAAATAATATATATTCCTCGGACAAAGAAATAGAGAAGTCAATCTCTGGGA 1680


```
QY 3841 ATCCATCCTTCTCAGCTTCAGCGCGGTATACCAACCCACCCCAATGCGCCACACTACT 3900
DB |||||||
3841 ATCCATCCTTCTCAGCTTCAGCGCGGTATACCAACCCACCCCAATGCGCCACACTACT 3900
QY |||||||
DB |||||||
3901 CAGGCTCATGTACAGTACAGGAATGTTCTTCTCATCCAACTGCGCCATGCGCCATGATG 3960
DB |||||||
3901 CAGGCTCATGTACAGTACAGGAATGTTCTTCTCATCCAACTGCGCCATGCGCCATGATG 3960
QY |||||||
DB |||||||
3961 CTAATGACGACACAGCCACCGCGCGGTCCCAAGGCGCGCTCGCTCAAAGTGCACTACAG 4020
DB |||||||
3961 CTAATGACGACACAGCCACCGCGCGGTCCCAAGGCGCGCTCGCTCAAAGTGCACTACAG 4020
QY |||||||
DB |||||||
4021 CCCATTCCAGTCTCAGCAACAGCGCATTTCCCTATATGACGACCCCTTCAGTCAAGGCC 4080
DB |||||||
4021 CCCATTCCAGTCTCAGCAACAGCGCATTTCCCTATATGACGACCCCTTCAGTCAAGGCC 4080
QY |||||||
DB |||||||
4081 CACCAACACAGCAGTGTGAAGCTGCGCTGGAGAACCGAAGGCCAAATTCCTCCTC 4140
DB |||||||
4081 CACCAACACAGCAGTGTGAAGCTGCGCTGGAGAACCGAAGGCCAAATTCCTCCTC 4140
QY |||||||
DB |||||||
4141 CTTCTACTGCTTCTACCAACTGGAAGCACAGAAACTAGAAATTCATTTATTTGTTTT 4200
DB |||||||
4141 CTTCTACTGCTTCTACCAACTGGAAGCACAGAAACTAGAAATTCATTTATTTGTTTT 4200
QY |||||||
DB |||||||
4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTGTTGAG 4260
DB |||||||
4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTGTTGAG 4260
QY |||||||
DB |||||||
4261 TGGAGATACCTTGGACCGAGTAGGCACTTAGGAACCTTGGGGCTATTCATATTTCCA 4320
DB |||||||
4261 TGGAGATACCTTGGACCGAGTAGGCACTTAGGAACCTTGGGGCTATTCATATTTCCA 4320
QY |||||||
DB |||||||
4321 TATGCTGTTTACAGTCCCGCAGGTACCCAGCTCTGTTGCCGAAACTGGAAGTTATTT 4380
DB |||||||
4321 TATGCTGTTTACAGTCCCGCAGGTACCCAGCTCTGTTGCCGAAACTGGAAGTTATTT 4380
QY |||||||
DB |||||||
4381 ATTTTAAATACCTTGAAGTCAATGAACTAGAACATCAGCTAGCAAAAGAAAGTAAAG 4440
DB |||||||
4381 ATTTTAAATACCTTGAAGTCAATGAACTAGAACATCAGCTAGCAAAAGAAAGTAAAG 4440
QY |||||||
DB |||||||
4441 GATTTCTGCTGCTATTTACTGCTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 4481
DB |||||||
4441 GATTTCTGCTGCTATTTACTGCTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 4481
```

RESULT 3

```
US-09-707-919A-20
; Sequence 20, Application US/09707919A
; Patent No. 6623927
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: Method for detection of human spinocerebellar ataxia 2
; FILE OF INVENTION: gene variants
; FILE REFERENCE: US 443
; CURRENT APPLICATION NUMBER: US/09/707,919A
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-707-919A-20
```

```
Query Match 100.0%; Score 4481; DB 3; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCGAGAAAGCAACCCAGCGCGCGCTCTCTACAGTGTCCCTCCCGGCCCGGG 60
DB |||||||
1 ACCCCGAGAAAGCAACCCAGCGCGCGCTCTCTACAGTGTCCCTCCCGGCCCGGG 60
```

```
QY 61 GCCACCTCAGCTTCTGCTTCCGTTCTGACCCCTCCGACTTCCGTTAAAGAGTCCCTATCCG 120
DB |||||||
61 GCCACCTCAGCTTCTGCTTCCGTTCTGACCCCTCCGACTTCCGTTAAAGAGTCCCTATCCG 120
QY |||||||
DB |||||||
121 CACCTTCGCTCCCAACCGCGCGCTCGGCGCGCCCTCCGATGCGGTCAAGCGCCGCA 180
DB |||||||
121 CACCTTCGCTCCCAACCGCGCGCTCGGCGCGCCCTCCGATGCGGTCAAGCGCCGCA 180
QY |||||||
DB |||||||
181 GCTCTCGAGTCCCGGCTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGAGTGGGCC 240
DB |||||||
181 GCTCTCGAGTCCCGGCTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGAGTGGGCC 240
QY |||||||
DB |||||||
241 GGGTGGCGCTCGCTCCAGCGCGCGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCG 300
DB |||||||
241 GGGTGGCGCTCGCTCCAGCGCGCGCGGCGGAGCGGCGGCGGCGGCGGCGGCGGCG 300
QY |||||||
DB |||||||
301 CCGGACCGTATCCTCCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB |||||||
301 CCGGACCGTATCCTCCCGCGCGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY |||||||
DB |||||||
361 CAGAGCTCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAACGCGCAGCGCGCGCGCG 420
DB |||||||
361 CAGAGCTCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAACGCGCAGCGCGCGCGCG 420
QY |||||||
DB |||||||
421 TTTTGGCGCGCGCTCCCGCGCGCTCTTGGTCTCGGCGGCGCTCCCGCGCGCTTCTGTCGTC 480
DB |||||||
421 TTTTGGCGCGCGCTCCCGCGCGCTCTTGGTCTCGGCGGCGCTCCCGCGCGCTTCTGTCGTC 480
QY |||||||
DB |||||||
481 GTCTTCTTCTCCCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB |||||||
481 GTCTTCTTCTCCCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY |||||||
DB |||||||
541 CTCGCGCGCGCTGCGCTCCCGCGCGCTTCCGCGCGCTCTCTTGGCGCGCGCGCGCG 600
DB |||||||
541 CTCGCGCGCGCTGCGCTCCCGCGCGCTTCCGCGCGCTCTCTTGGCGCGCGCGCGCG 600
QY |||||||
DB |||||||
601 TGTCCCGCGCGCTGCGAGCGGTGTATGCGGCGCTTCAACATGTCTGCTGAAGCGCGCG 660
DB |||||||
601 TGTCCCGCGCGCTGCGAGCGGTGTATGCGGCGCTTCAACATGTCTGCTGAAGCGCGCG 660
QY |||||||
DB |||||||
661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
DB |||||||
661 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
QY |||||||
DB |||||||
721 CAGCGCGCGCGCGCGCTGCGCAATGTCTCGCAAGCGCGCGCGCGCGCGCGCGCGCG 780
DB |||||||
721 CAGCGCGCGCGCGCGCTGCGCAATGTCTCGCAAGCGCGCGCGCGCGCGCGCGCGCG 780
QY |||||||
DB |||||||
781 CCGCGCGCGCGCGCTTCCCGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 840
DB |||||||
781 CCGCGCGCGCGCGCTTCCCGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 840
QY |||||||
DB |||||||
841 TCCTCGGTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB |||||||
841 TCCTCGGTGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY |||||||
DB |||||||
901 AGTAACAAAGGAGTCTGCTACGATTTCTTTTGTGTAAGTCTATGCAAAATATGAGG 960
DB |||||||
901 AGTAACAAAGGAGTCTGCTACGATTTCTTTTGTGTAAGTCTATGCAAAATATGAGG 960
QY |||||||
DB |||||||
961 ATGTTTCATATATCTTACATCATGTTTCTGCTCCAAATGTGAAGTCAAGTGAAGTGA 1020
DB |||||||
961 ATGTTTCATATATCTTACATCATGTTTCTGCTCCAAATGTGAAGTCAAGTGAAGTGA 1020
QY |||||||
DB |||||||
1021 GGTATATATGAGGAGTATTTTAAACCTTACAGTCCGAGTGTGATTTTGGTCTTCTATGCC 1080
DB |||||||
1021 GGTATATATGAGGAGTATTTTAAACCTTACAGTCCGAGTGTGATTTTGGTCTTCTATGCC 1080
QY |||||||
DB |||||||
1081 GCACATGAGAAAGTACAGAAATCCAGTTTCGGGCGCGCGCGCGCGCGCGCGCGCG 1140
DB |||||||
1081 GCACATGAGAAAGTACAGAAATCCAGTTTCGGGCGCGCGCGCGCGCGCGCGCGCG 1140
QY |||||||
DB |||||||
1141 ATTTTGTTCAAATGTTTCAGACTTTTGTGTGTAAGTCTTAAAGTATGAGTCCAGTTAT 1200
```

|||||
1141 ATTTTGTTCANAATGTCAGACCTTTGTTGGTACAGTTTAAAGATATGACCTCCAGTTAT 1200
QY
1201 GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTGAATGGCGAAACACAAA 1260
Db
1201 GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAAGTGAATGGCGAAACACAAA 1260
QY
1261 GAGAAGGACCTGGAGCCCTGGGATGAGGTGAACTCACAGCCAAATGAGGAACTTGAAGGCT 1320
Db
1261 GAGAAGGACCTGGAGCCCTGGGATGAGGTGAACTCACAGCCAAATGAGGAACTTGAAGGCT 1320
QY
1321 TTGGAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380
Db
1321 TTGGAAAATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380
QY
1381 AATTATGTTGATGTTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGGCCCTTAGAA 1440
Db
1381 AATTATGTTGATGTTCTACGTATGATAGCAGTTTATCTTCGTATACAGTGGCCCTTAGAA 1440
QY
1441 AGAGATAAATCAGAGAAATTTTAAACCGGGAAGCAAGGGCAAAACAGTTAGCAGAGAA 1500
Db
1441 AGAGATAAATCAGAGAAATTTTAAACCGGGAAGCAAGGGCAAAACAGTTAGCAGAGAA 1500
QY
1501 ATTTAGTCAAGTGCCTCAGTACAAAGCTCGAGTGGCCCTGGAATGATATAGGAGTGAG 1560
Db
1501 ATTTAGTCAAGTGCCTCAGTACAAAGCTCGAGTGGCCCTGGAATGATATAGGAGTGAG 1560
QY
1561 GAAGAAAATACACAGCAGTTACAGAGAAATTCAGTGAACCTGAGGGGACACAGCATAAATC 1620
Db
1561 GAAGAAAATACACAGCAGTTACAGAGAAATTCAGTGAACCTGAGGGGACACAGCATAAATC 1620
QY
1621 ACTAGGGAAAATAAATATATATTCCTCTGGACAAAGAAATAGAGAAATCATATCTTGGGA 1680
Db
1621 ACTAGGGAAAATAAATATATATTCCTCTGGACAAAGAAATAGAGAAATCATATCTTGGGA 1680
QY
1681 AGTGGAGACAGAAATTCACCGGCTATGGGCGAGCCCTGGATCGGGCTCCATGCCATCAAG 1740
Db
1681 AGTGGAGACAGAAATTCACCGGCTATGGGCGAGCCCTGGATCGGGCTCCATGCCATCAAG 1740
QY
1741 TCCACTTCTCACACTTCAGATTTCAACCGAAATTCCTGGTTTCAGACCAAGAGATGATTAAT 1800
Db
1741 TCCACTTCTCACACTTCAGATTTCAACCGAAATTCCTGGTTTCAGACCAAGAGATGATTAAT 1800
QY
1801 GGAAGTGTTCCTCGGCATCGCCTTGCCCATCTCTCTCTCTCGCCACCTTCTCGCTAC 1860
Db
1801 GGAAGTGTTCCTCGGCATCGCCTTGCCCATCTCTCTCTCTCGCCACCTTCTCGCTAC 1860
QY
1861 CAGTCAGTGCCTCAACTCTCTTCCACCTCGGGCAGCCACCCCTACACGGCGGCCCTCAGG 1920
Db
1861 CAGTCAGTGCCTCAACTCTCTTCCACCTCGGGCAGCCACCCCTACACGGCGGCCCTCAGG 1920
QY
1921 CCCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTCTCATGTTCTCCAGCTCCT 1980
Db
1921 CCCCCCTCGGGCCATCCAGACCCCGTCTCACCCCTCTCTCATGTTCTCCAGCTCCT 1980
QY
1981 GTCTCTACTATGCTTAACGCAATGCTTCTCAGAGAGGCTCTCAAGGATGTCCCAAAGGCC 2040
Db
1981 GTCTCTACTATGCTTAACGCAATGCTTCTCAGAGAGGCTCTCAAGGATGTCCCAAAGGCC 2040
QY
2041 CAGCGACATCCTCGAAATCAACAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100
Db
2041 CAGCGACATCCTCGAAATCAACAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100
QY
2101 GAATTTGTTATCCCAACACCCAGTGAAGAGCTACTCTCTCAGTGAAGAGCCAGT 2160
Db
2101 GAATTTGTTATCCCAACACCCAGTGAAGAGCTACTCTCTCAGTGAAGAGCCAGT 2160
QY
2161 CCCTCGGGGGAAAGTGGTTCATCAGTGGTCAAGTGGGTTTCAAGATTTATCCCTTAAACT 2220
Db
2161 CCCTCGGGGGAAAGTGGTTCATCAGTGGTCAAGTGGGTTTCAAGATTTATCCCTTAAACT 2220
QY
2221 CATAGACCGAGGTCTCCGACAGACAGATTTGGAAAATACCCCCAGTGGGCCAGTTCTT 2280

Db
2221 CATAGACCGAGGTCTCCGACAGACAGATTTGGAAAATACCCCCAGTGGGCCAGTTCTT 2280
QY
2281 GCTTCTCCCAAGCTGTGTTATTTCCAACTGAAGCTGTGTGCCATGCTTATTCAGCTGCA 2340
Db
2281 GCTTCTCCCAAGCTGTGTTATTTCCAACTGAAGCTGTGTGCCATGCTTATTCAGCTGCA 2340
QY
2341 TCTCTCAGCCTGCTAGTCTGCTGCAACAGAGCTGTATACCCCTTCTAGTCAGGCTAAA 2400
Db
2341 TCTCTCAGCCTGCTAGTCTGCTGCAACAGAGCTGTATACCCCTTCTAGTCAGGCTAAA 2400
QY
2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTGAGGGGAATAAAGAAAATATATAA 2460
Db
2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTGAGGGGAATAAAGAAAATATATAA 2460
QY
2461 CCCAATGAACATCATCCTAGTCTTCTCAAAAGCTGAAAAAAGGATATATCAACAGTTGTT 2520
Db
2461 CCCAATGAACATCATCCTAGTCTTCTCAAAAGCTGAAAAAAGGATATATCAACAGTTGTT 2520
QY
2521 TCTGAAACATAGAAAACAGATTTGATGATTTAAAGAAAATTTAAGAAATGATTTTAGTTACAG 2580
Db
2521 TCTGAAACATAGAAAACAGATTTGATGATTTAAAGAAAATTTAAGAAATGATTTTAGTTACAG 2580
QY
2581 CCAAGTCTTACTTCTGAAATCTATGGATCAACTACTATAAACAAGGATAGAGAGGAGAAAAA 2640
Db
2581 CCAAGTCTTACTTCTGAAATCTATGGATCAACTACTATAAACAAGGATAGAGAGGAGAAAAA 2640
QY
2641 TCAAGAGATTTTGTATCAAAAGACAAAATTTGAACCAAGTGTCTAAGGATTTCTTCAATGAAAAAT 2700
Db
2641 TCAAGAGATTTTGTATCAAAAGACAAAATTTGAACCAAGTGTCTAAGGATTTCTTCAATGAAAAAT 2700
QY
2701 AGCAGCAGCAACTGTACAGTGGGAGCAGCAAGCCGAATAGCCCGAGCAATTTCCCTTCA 2760
Db
2701 AGCAGCAGCAACTGTACAGTGGGAGCAGCAAGCCGAATAGCCCGAGCAATTTCCCTTCA 2760
QY
2761 ATACTTAGTAAACACGGAGCACAAGAGGGGACCTGAGTCACTTCCCAAGGGGTTTCAAGACT 2820
Db
2761 ATACTTAGTAAACACGGAGCACAAGAGGGGACCTGAGTCACTTCCCAAGGGGTTTCAAGACT 2820
QY
2821 TCCAGCAGCAGTGTAAACAAGAGAAAGACGTAAGGAAAGAGAAAGACGAGCTGAG 2880
Db
2821 TCCAGCAGCAGTGTAAACAAGAGAAAGACGTAAGGAAAGAGAAAGACGAGCTGAG 2880
QY
2881 CAAAGTGAAGAAATCAACATTTGAATCCCAATGCAAGAGGATTCACCCAGTGTCTCTCT 2940
Db
2881 CAAAGTGAAGAAATCAACATTTGAATCCCAATGCAAGAGGATTCACCCAGTGTCTCTCT 2940
QY
2941 CAGCCAAAGCCCTTCTACTACCCCACTTCCACCTCGGCTCAAGACAACTAGCCCATCT 3000
Db
2941 CAGCCAAAGCCCTTCTACTACCCCACTTCCACCTCGGCTCAAGACAACTAGCCCATCT 3000
QY
3001 ATGGTGGGTCTATCAACAGCCTCACTCCAGTTTATCTCAGCCTGTGTTTGTGACCAAAAT 3060
Db
3001 ATGGTGGGTCTATCAACAGCCTCACTCCAGTTTATCTCAGCCTGTGTTTGTGACCAAAAT 3060
QY
3061 ATGATGTATCCAGTCCAGTGGAGCCAGCGGTGCAACCTTTATACCAATACCTATGACG 3120
Db
3061 ATGATGTATCCAGTCCAGTGGAGCCAGCGGTGCAACCTTTATACCAATACCTATGACG 3120
QY
3121 CCCATGCCAGTGAATCAAGCCCAAGCATATAGAGCAGTACCAAAATATGCCCAAGAGCGG 3180
Db
3121 CCCATGCCAGTGAATCAAGCCCAAGCATATAGAGCAGTACCAAAATATGCCCAAGAGCGG 3180
QY
3181 CAAAGCAGCAGTATCAGAGTGCATGATGACCCAGCGTTCAGCAGCGGGCCACCGAT 3240
Db
3181 CAAAGCAGCAGTATCAGAGTGCATGATGACCCAGCGTTCAGCAGCGGGCCACCGAT 3240
QY
3241 GAGCCACCCACAGCTTACTCCACGCAATATGTTGGCTACAGTCTCAGAGTTCCCA 3300
Db
3241 GAGCCACCCACAGCTTACTCCACGCAATATGTTGGCTACAGTCTCAGAGTTCCCA 3300
QY
3301 AATCAGCCCCCTTGTTCAGCATGTGCCACATTTACGTCTCAGCATCTCATGTCTATAGT 3360
Db
3301 AATCAGCCCCCTTGTTCAGCATGTGCCACATTTACGTCTCAGCATCTCATGTCTATAGT 3360

QY 3361 CCTGTAATACAGGGTAATGCTAGATGATGGACACCAACACAGCCCGCCTGGTTTA 3420
Db 3361 CCTGTAATACAGGGTAATGCTAGATGATGGACACCAACACAGCCCGCCTGGTTTA 3420
QY 3421 GTATCTTTCTTACGAACTCAGTACGGGGCTATGAGCAGACGATGCGATGATCATGT 3480
Db 3421 GTATCTTTCTTACGAACTCAGTACGGGGCTATGAGCAGACGATGCGATGATCATGT 3480
QY 3481 CCCAAATTACCATCAACAAGGAGACAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3540
Db 3481 CCCAAATTACCATCAACAAGGAGACAAGCCCTTCTTCTACTTTGGCCATTTCCACGGGC 3540
QY 3541 TCCCTTGTCTACGAGTATGCGCACCTTAACGCTACCTTGCCACCCACATATCTCCACACCT 3600
Db 3541 TCCCTTGTCTACGAGTATGCGCACCTTAACGCTACCTTGCCACCCACATATCTCCACACCT 3600
QY 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGGCAACATGTTGGAGTCACTCTGCA 3660
Db 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGGCAACATGTTGGAGTCACTCTGCA 3660
QY 3661 CCCAGTCTCTTTCAGCACCATCAGCACAGGCGCCGAGGCTCTCCATCTGSCCAGTCCA 3720
Db 3661 CCCAGTCTCTTTCAGCACCATCAGCACAGGCGCCGAGGCTCTCCATCTGSCCAGTCCA 3720
QY 3721 CAGCAGCAGTACGCCATTTACCACGCGGGCTTGGCCCAACTCCACCTCCATGACACT 3780
Db 3721 CAGCAGCAGTACGCCATTTACCACGCGGGCTTGGCCCAACTCCACCTCCATGACACT 3780
QY 3781 GCCTCCACAGCAGTCCGACAGATAGTTTCCAGCAGCAACAGACTGCTTTAGG 3840
Db 3781 GCCTCCACAGCAGTCCGACAGATAGTTTCCAGCAGCACAACAGACTGCTTTAGG 3840
QY 3841 ATCCATCTCTTCAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCACCTACT 3900
Db 3841 ATCCATCTCTTCAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCACCTACT 3900
QY 3901 CAGGCTCATGTACAGTCAAGATGTTCTTCTCATCAACTGCCCACTGCCCAATGATG 3960
Db 3901 CAGGCTCATGTACAGTCAAGATGTTCTTCTCATCAACTGCCCACTGCCCAATGATG 3960
QY 3961 CTAATGACGACACAGCCACCGCGGTCCAGGCGCGCTCGCTCAAAGTGCACTAG 4020
Db 3961 CTAATGACGACACAGCCACCGCGGTCCAGGCGCGCTCGCTCAAAGTGCACTAG 4020
QY 4021 CCATTTCCAGTCTCGACAAACAGCGCATTTCCCTATATGACGACCCCTTCAGTACAAGC 4080
Db 4021 CCATTTCCAGTCTCGACAAACAGCGCATTTCCCTATATGACGACCCCTTCAGTACAAGC 4080
QY 4081 CACCACCAACAGCAGTGTGTAAGGCTGCCCTGAGGAAACCGAAAGGCCAAATTTCCCTCTC 4140
Db 4081 CACCACCAACAGCAGTGTGTAAGGCTGCCCTGAGGAAACCGAAAGGCCAAATTTCCCTCTC 4140
QY 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGACAGAACTAGAAATTTCAATTTATTTT 4200
Db 4141 CTTTCTACTGCTTCTACCAACTGGAAGCAGACAGAACTAGAAATTTCAATTTATTTT 4200
QY 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTTCAGTTCAG 4260
Db 4201 TAAATATATATGTTGATTTCTTTGTAACATCAATAGGAATGCTAACAGTTTCAGTTCAG 4260
QY 4261 TGGAGATACCTTGGACCGAGTAGGCAATTTAGGAACCTTGGGGCTATTCCATATTTCCA 4320
Db 4261 TGGAGATACCTTGGACCGAGTAGGCAATTTAGGAACCTTGGGGCTATTCCATATTTCCA 4320
QY 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGAGTTATTT 4380
Db 4321 TATGCTGTTTCAGAGTCCGCGAGGTACCCAGCTCTGCTTGCAGAACTGGAGTTATTT 4380
QY 4381 ATTTTAAATACCCCTTGAAAGTCAATGAACATCAGCTAGCAAAAGAGTAACAGAGT 4440
Db 4381 ATTTTAAATACCCCTTGAAAGTCAATGAACATCAGCTAGCAAAAGAGTAACAGAGT 4440

QY 4441 GATTCTTGCTCTACTTACTGCTCTAAAAA 4481
Db 4441 GATTCTTGCTCTACTTACTGCTCTAAAAA 4481

RESULT 4
US-09-083-268-2
; Sequence 2, Application US/09083268
; Patent No. 6673535
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
; TITLE OF INVENTION: ATAXIA-2 AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Muetling, Raasch & Gebhardt, P.A.
; STREET: 119 No. 6673535th Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/727,084
; FILING DATE: 08-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McCormack, Myra H.
; REGISTRATION NUMBER: 36,602
; REFERENCE/DOCKET NUMBER: 232.00010101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/305-1220
; TELEFAX: 612/305-1228
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4481 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 163..4101
US-09-083-268-2

Query Match 100.0%; Score 4481; DB 3; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCCCGAAGAACCAACCCAGCGCGCCGCTCCCTCACGTGCTCCCTCCCGGCCCCGGG 60
Db 1 ACCCCCGAAGAACCAACCCAGCGCGCGCCGCTCCCTCACGTGCTCCCTCCCGGCCCCGGG 60
QY 61 GCCACCTCACTGCTTCTGCTTCCGTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCG 120
Db 61 GCCACCTCACTGCTTCTGCTTCCGTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCG 120
QY 121 CACCTCCGCTCCCAACCCGCGCGCTCCGCGCGCCCTCCGATCGCTCAGCGCGCGCA 180
Db 121 CACCTCCGCTCCCAACCCGCGCGCTCCGCGCGCCCTCCGATCGCTCAGCGCGCGCA 180
QY 181 GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCCGCGCAGCAGAGTGGGCC 240
Db 181 GCTCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCCGCGCAGCAGAGTGGGCC 240
QY 241 GGGTGGCGCTCGCTTCCAGCGCGCGCGCGGAGCGGCGCGGTGGCGCGGCC 300

QY 2461 CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCAACAGTTGTT 2520
Db |||||
2461 CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCAACAGTTGTT 2520
QY |||||
Db |||||
2521 TCTGAAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAAATGATTTAGGTTACAG 2580
Db |||||
2521 TCTGAAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAGAAATGATTTAGGTTACAG 2580
QY |||||
Db |||||
2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTATCAAAACAAAAATAGAGAGGAGAAAAA 2640
Db |||||
2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTATCAAAACAAAAATAGAGAGGAGAAAAA 2640
QY |||||
Db |||||
2641 TCAAGAGATTTGATCAAAAGAChAAATTTGAAACCAAGTGTCTAGGATTTCTTCAATGAAAT 2700
Db |||||
2641 TCAAGAGATTTGATCAAAAGACAAAAATTGAACCAAGTGTCTAGGATTTCTTCAATGAAAT 2700
QY |||||
Db |||||
2701 AGCAGCAGCAACTGTACAGTGGCAGCAGCAGCGGAATAGCCCAAGCATTTCCCTTCA 2760
Db |||||
2701 AGCAGCAGCAACTGTACAGTGGCAGCAGCAGCGGAATAGCCCAAGCATTTCCCTTCA 2760
QY |||||
Db |||||
2761 ATACTTTAGTAAACACGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCAACT 2820
Db |||||
2761 ATACTTTAGTAAACACGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCAACT 2820
QY |||||
Db |||||
2821 TCCAGCCCGCATGTAAACAGAGAAAGACGATAGGAAGAGAAAGACGAGCTGAG 2880
Db |||||
2821 TCCAGCCCGCATGTAAACAGAGAAAGACGATAGGAAGAGAAAGACGAGCTGAG 2880
QY |||||
Db |||||
2881 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTCTCT 2940
Db |||||
2881 CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGAGTTCAACCCAGCTTCTCTCT 2940
QY |||||
Db |||||
2941 CAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAAGCACAACCTAGGCCATCT 3000
Db |||||
2941 CAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAAGCACAACCTAGGCCATCT 3000
QY |||||
Db |||||
3001 ATGGTGGGTCAATCAACAGCAAACTCCAGTTTATATCTCAGCCCTGTTTGTGACCAAAAT 3060
Db |||||
3001 ATGGTGGGTCAATCAACAGCAAACTCCAGTTTATATCTCAGCCCTGTTTGTGACCAAAAT 3060
QY |||||
Db |||||
3061 ATGATGTATCCAGTCCAGTCCAGCCAGCGGTGCAACCTTTATACCAATACCTATGAG 3120
Db |||||
3061 ATGATGTATCCAGTCCAGTCCAGCCAGCGGTGCAACCTTTATACCAATACCTATGAG 3120
QY |||||
Db |||||
3121 CCATGCCAGTGAATCAAGCCAAAGACATATAGCAGCAGTACCAATATGCCCAACAGCGG 3180
Db |||||
3121 CCATGCCAGTGAATCAAGCCAAAGACATATAGCAGCAGTACCAATATGCCCAACAGCGG 3180
QY |||||
Db |||||
3181 CAAAGCCAGCATCATAGTCCCATATGCAACCCAGCGTCAGCAGCGGGCCACCGAT 3240
Db |||||
3181 CAAAGCCAGCATCATAGTCCCATATGCAACCCAGCGTCAGCAGCGGGCCACCGAT 3240
QY |||||
Db |||||
3241 GCAGCAGCCCAACAGCTTACTCCAGCCAAATATGTTGCCCTACAGTCCCTCAGCAGTTCCCA 3300
Db |||||
3241 GCAGCAGCCCAACAGCTTACTCCAGCCAAATATGTTGCCCTACAGTCCCTCAGCAGTTCCCA 3300
QY |||||
Db |||||
3301 AATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCCTCATGTCTATAGT 3360
Db |||||
3301 AATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCCTCATGTCTATAGT 3360
QY |||||
Db |||||
3361 CTTGTAATACAGGGTAATGCTAGATGATGGCACCAATATCAGTCTCAGCATCCTCATGTCTATAGT 3420
Db |||||
3361 CTTGTAATACAGGGTAATGCTAGATGATGGCACCAATATCAGTCTCAGCATCCTCATGTCTATAGT 3420
QY |||||
Db |||||
3421 GTATCTTTCTTACGCAACTCAGTACGGGGCTCATGAGCAGCATGCGATGATGCACTGT 3480
Db |||||
3421 GTATCTTTCTTACGCAACTCAGTACGGGGCTCATGAGCAGCATGCGATGATGCACTGT 3480
QY |||||
Db |||||
3481 CCCAAATTACCATCAACAAAGGAGACAAGCCCTTCTTCTACTTTGCCATTTCCACGGGC 3540
Db |||||
3481 CCCAAATTACCATCAACAAAGGAGACAAGCCCTTCTTCTACTTTGCCATTTCCACGGGC 3540

QY 3541 TCCCTTGTCTAGCAGTATGCGCACCCCTAAACGCTACCTTGCAACCACATACTTCCACACCTT 3600
Db |||||
3541 TCCCTTGTCTAGCAGTATGCGCACCCCTAAACGCTACCTTGCAACCACATACTTCCACACCTT 3600
QY |||||
Db |||||
3601 CAGCCTTCAGCTACCCCTACCTGGACAGCAGCAAGCCAAACATGTTGGAAGTCATCTCTGCA 3660
Db |||||
3601 CAGCCTTCAGCTACCCCTACCTGGACAGCAGCAAGCCAAACATGTTGGAAGTCATCTCTGCA 3660
QY |||||
Db |||||
3661 CCCAGTCTCTGTTTACAGCACCATCAGCAGCAGCCGCCAGGCTCTCCATCTGGCCAGTCCA 3720
Db |||||
3661 CCCAGTCTCTGTTTACAGCACCATCAGCAGCAGCCGCCAGGCTCTCCATCTGGCCAGTCCA 3720
QY |||||
Db |||||
3721 CAGCAGCAGTCAAGCATTTTACACGCGGGGCTTGGCCCAACTCCACCTCTCATGACACT 3780
Db |||||
3721 CAGCAGCAGTCAAGCATTTTACACGCGGGGCTTGGCCCAACTCCACCTCTCATGACACT 3780
QY |||||
Db |||||
3781 GCCTCCAAACGCGAGTCGCCACAGAAATAGTTTCCACGAGCAGCACAAGACTGTCTTTACG 3840
Db |||||
3781 GCCTCCAAACGCGAGTCGCCACAGAAATAGTTTCCACGAGCAGCACAAGACTGTCTTTACG 3840
QY |||||
Db |||||
3841 ATCCATCCTTCTCAGCTTTCAGCCGCGGTATACCAACCCACACATGGGCCACGATACCT 3900
Db |||||
3841 ATCCATCCTTCTCAGCTTTCAGCCGCGGTATACCAACCCACACATGGGCCACGATACCT 3900
QY |||||
Db |||||
3901 CAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCCAACTGCCCATGCGCCCAATGATG 3960
Db |||||
3901 CAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCCAACTGCCCATGCGCCCAATGATG 3960
QY |||||
Db |||||
3961 CTAATGAGCAGACAGCCACCCGCGGTCCCGAGCGCCCTCGCTCAAGAGTGCACTACAG 4020
Db |||||
3961 CTAATGAGCAGACAGCCACCCGCGGTCCCGAGCGCCCTCGCTCAAGAGTGCACTACAG 4020
QY |||||
Db |||||
4021 CCCATTCCAGTCTCGAACAACAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAAGCC 4080
Db |||||
4021 CCCATTCCAGTCTCGAACAACAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAAGCC 4080
QY |||||
Db |||||
4081 CACCACCAACAGCAGTGTAAAGCTGCGCTGGAGGAACCGAAAGGCCAAATTTCCCTCTCTC 4140
Db |||||
4081 CACCACCAACAGCAGTGTAAAGCTGCGCTGGAGGAACCGAAAGGCCAAATTTCCCTCTCTC 4140
QY |||||
Db |||||
4141 CTTTCTACTGCTTCTACCACTGGAAGCAGCAAACTAGAAATTTTATTTTGTGTTT 4200
Db |||||
4141 CTTTCTACTGCTTCTACCACTGGAAGCAGCAAACTAGAAATTTTATTTTGTGTTT 4200
QY |||||
Db |||||
4201 TAAATATATATGTTGATTTCTTGTAAACATCCAATAGGAATGCTAAACAGTTTCACTTGCAG 4260
Db |||||
4201 TAAATATATATGTTGATTTCTTGTAACTCAATAGGAATGCTAAACAGTTTCACTTGCAG 4260
QY |||||
Db |||||
4261 TGGAGATACCTTGGACCGAGTAGAGGCAATTTAGGAATTTGGGGCTATTCCATTAATTCCA 4320
Db |||||
4261 TGGAGATACCTTGGACCGAGTAGAGGCAATTTAGGAATTTGGGGCTATTCCATAATTCCA 4320
QY |||||
Db |||||
4321 TATGCTGTTTTCAGAGTCCCGAGGTACCCAGCTCTGCTTCCGAAATCGGAAGTTATTT 4380
Db |||||
4321 TATGCTGTTTTCAGAGTCCCGAGGTACCCAGCTCTGCTTCCGAAATCGGAAGTTATTT 4380
QY |||||
Db |||||
4381 ATTTTTTAATACCTTTGAAAGTCATGAACAATCAGCTAGCAAAAAGTAACAAGAGT 4440
Db |||||
4381 ATTTTTTAATACCTTTGAAAGTCATGAACAATCAGCTAGCAAAAAGTAACAAGAGT 4440
QY |||||
Db |||||
4441 GATTCTTGTGCTATTAATGCTATAAAAAA 4481
Db |||||
4441 GATTCTTGTGCTATTAATGCTATAAAAAA 4481

RESULT 5

US-08-981-998A-2

; Sequence 2, Application US/08981998A

; Patent No. 6844431

; GENERAL INFORMATION:

; APPLICANT: PULST, STEFAN M.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR

; ATAXIA-2 AND PRODUCTS RELATED THERETO

NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSER: MUEITING, RAASCH & GEBHARDT, P.A.
STREET: 119 NORTH FOURTH STREET, SUITE 203
CITY: MINNEAPOLIS
STATE: MINNESOTA
COUNTRY: 55401
ZIP: 55401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: PatentIn
FILING DATE: 11-May-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 97/42314
FILING DATE: 08-MAY-1997
APPLICATION NUMBER: US 08/727,084
FILING DATE: 08-OCT-1996
APPLICATION NUMBER: US 60/022,207
FILING DATE: 19-JUL-1996
APPLICATION NUMBER: US 60/017,388
FILING DATE: 08-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: MUEITING, ANN M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 232.00010120
TELEPHONE: 612-305-1217
TELEFAX: 612-305-1228
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4481 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 163..4101
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-981-998A-2

Query Match 100.0%; Score 4481; DB 3; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCCCGAGAAAGCAACCCAGCGCGCGCTCCTCAGCTGTCCCTCCCGGCCCGGG 60
DB 1 ACCCCGAGAAAGCAACCCAGCGCGCGCTCCTCAGCTGTCCCTCCCGGCCCGGG 60
QY 61 GCACCTCAGCTTCTGCTCCGCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
DB 61 GCACCTCAGCTTCTGCTCCGCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
QY 121 CACCTCCGCTCCACCCGCGCTCCGCGCGCGCGCTCCGATCGCTCAGCGCGCGCA 180
DB 121 CACCTCCGCTCCACCCGCGCTCCGCGCGCGCGCTCCGATCGCTCAGCGCGCGCA 180
QY 181 GCTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCCGCGCAGCGGTGGCCC 240
DB 181 GCTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCCGCGCAGCGGTGGCCC 240
QY 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 GGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 CCGGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 CCGGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360

QY 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGTAGCAACGGCAACGGCGCGCGCGG 420
DB 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGTAGCAACGGCAACGGCGCGCGCGG 420
QY 421 TTTCCGCGCGCGCTCCGCGCGCTCCTTGGTCTCGGCGCGCGCTCCCGCGCGCTTCTGCTGC 480
DB 421 TTTCCGCGCGCGCTCCGCGCGCTCCTTGGTCTCGGCGCGCGCTCCCGCGCGCTTCTGCTGC 480
QY 481 GTCTCTTCTCCCTCCGCTCAGCGCGCGCTCCGCGCGCGCTCCGCGCGCGCTCCCGG 540
DB 481 GTCTCTTCTCCCTCCGCTCAGCGCGCGCTCCGCGCGCGCTCCGCGCGCGCTCCCGG 540
QY 541 CTCGCGCGCGCTCGCTCCCGCGCGCTTCCGCGCGCTTCTTGGCGCGCGCGCTCCCGG 600
DB 541 CTCGCGCGCGCTCGCTCCCGCGCGCTTCCGCGCGCTTCTTGGCGCGCGCGCTCCCGG 600
QY 601 TGTCCCGCGCGCTCGCGCGCTGATGGCGCGCTTCCGCGCGCTTCCGCGCGCGCTCCCGG 660
DB 601 TGTCCCGCGCGCTCGCGCGCTGATGGCGCGCTTCCGCGCGCTTCCGCGCGCGCTCCCGG 660
QY 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
DB 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
QY 721 CAGCGCGCGCGCTCGCTCCGCGCGCTTCCGCGCGCTTCCGCGCGCGCTTCTAGGCTCG 780
DB 721 CAGCGCGCGCGCTCGCTCCGCGCGCTTCCGCGCGCTTCCGCGCGCGCTTCTAGGCTCG 780
QY 781 CCGCGCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCGCTTCCG 840
DB 781 CCGCGCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCGCTTCCG 840
QY 841 TCCTCGCTGCTCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCT 900
DB 841 TCCTCGCTGCTCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCTTCCGCGCGCT 900
QY 901 AGTAAACAAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAATATGAGG 960
DB 901 AGTAAACAAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAATATGAGG 960
QY 961 ATGTTTCATATATACATCAGTGTGCTCCAAATGGAAGTCAAGTGAAGTGAAGTGA 1020
DB 961 ATGTTTCATATATACATCAGTGTGCTCCAAATGGAAGTCAAGTGAAGTGAAGTGA 1020
QY 1021 GGTATATATGAAGGAGTGTAAACCTTACAGTCCGAACTTACAGTCCGAACTTGTGCTGCT 1080
DB 1021 GGTATATATGAAGGAGTGTAAACCTTACAGTCCGAACTTACAGTCCGAACTTGTGCTGCT 1080
QY 1081 GCACATGAGAAAGTACAGAAATCCAGTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1081 GCACATGAGAAAGTACAGAAATCCAGTTCGCGCGCGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 1141 ATTTTGTTCAAATGTTTACAGTTCGCTATCTGCTATCTGCTTAAAGATATGAGTCCAGTAT 1200
DB 1141 ATTTTGTTCAAATGTTTACAGTTCGCTATCTGCTATCTGCTTAAAGATATGAGTCCAGTAT 1200
QY 1201 GCAAAAGAGATGCTTTTAACTGCTATCTGCTATCTGCTTAAAGATATGAGTCCAGTAT 1260
DB 1201 GCAAAAGAGATGCTTTTAACTGCTATCTGCTATCTGCTTAAAGATATGAGTCCAGTAT 1260
QY 1261 GAGAGGAGCTCGAGCGCTGGGATGAGTGAATCAGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
DB 1261 GAGAGGAGCTCGAGCGCTGGGATGAGTGAATCAGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
QY 1321 TTGGAAGATGAGCTATCTAATGGATGGGATCCCAATGATATGTTTCGATATATGAGAA 1380
DB 1321 TTGGAAGATGAGCTATCTAATGGATGGGATCCCAATGATATGTTTCGATATATGAGAA 1380
QY 1381 AATTATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCTGTTATACAGTGCCTTTAGAA 1440
DB 1381 AATTATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCTGTTATACAGTGCCTTTAGAA 1440

QY 1441 AGAGATTAAGTCTGAGAGAAATTTTAAACCGGGAAGCAAGGGGCAAAACAGTTAGCAGAAAGAA 1500
DB |||||
DB 1441 AGAGATAACTCTGAGAGAAATTTTAAACCGGGAAGCAAGGGGCAAAACAGTTAGCAGAAAGAA 1500
QY |||||
QY 1501 ATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTGGAAAAATGATGATAGGAGTGCAG 1560
DB |||||
DB 1501 ATTGAGTCAAGTGGCCAGTACAAAGCTCGAGTGGCCCTGGAAAAATGATGATAGGAGTGCAG 1560
QY |||||
QY 1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAAC 1620
DB |||||
DB 1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAAC 1620
QY |||||
QY 1621 ACTAGGGAAAAATAATATATATCTCTCTGGAACAAAGAAATAGAGAAAGTCAATATCTCTGGGGA 1680
DB |||||
DB 1621 ACTAGGGAAAAATAATATATATCTCTCTGGAACAAAGAAATAGAGAAAGTCAATATCTCTGGGGA 1680
QY |||||
QY 1681 AGTGGGAGACAGAAATTCACCGGCTATGGCCAGCCCTGAGTGGGCTCCATGCCATCAAGA 1740
DB |||||
DB 1681 AGTGGGAGACAGAAATTCACCGGCTATGGCCAGCCCTGAGTGGGCTCCATGCCATCAAGA 1740
QY |||||
QY 1741 TCCACTTCTCACACTTCAGATTTCAACCCCGAATTCCTGGTTTCAGACCAAGAGTAGTTAAAT 1800
DB |||||
DB 1741 TCCACTTCTCACACTTCAGATTTCAACCCCGAATTCCTGGTTTCAGACCAAGAGTAGTTAAAT 1800
QY |||||
QY 1801 GGAGTGTTCCTGGCCATCGCCTTGGCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
DB |||||
DB 1801 GGAGTGTTCCTGGCCATCGCCTTGGCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
QY |||||
QY 1861 CAGTCAGTTCACACT 1920
DB |||||
DB 1861 CAGTCAGTTCACACT 1920
QY |||||
QY 1921 CCCCCCTCGGGCCATCAGACCCCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGCTCCT 1980
DB |||||
DB 1921 CCCCCCTCGGGCCATCAGACCCCCCGTCTCACCCCTCTGCTCATGGTTCTCCAGCTCCT 1980
QY |||||
QY 1981 GTCTCTACTATGCCCTAAACGCAATGTTCTCAGAGGGCCCTCCAGGATGTCGCCAAGGCC 2040
DB |||||
DB 1981 GTCTCTACTATGCCCTAAACGCAATGTTCTCAGAGGGCCCTCCAGGATGTCGCCAAGGCC 2040
QY |||||
QY 2041 CAGCGACATCTCTCGAAATCACAGAGTTTCTGCTGGGAGGGTTCCATATCCAGTGGCCCTA 2100
DB |||||
DB 2041 CAGCGACATCTCTCGAAATCACAGAGTTTCTGCTGGGAGGGTTCCATATCCAGTGGCCCTA 2100
QY |||||
QY 2101 GAAATTTGATCCCAACCCAGTGAAGCAGCTACTCTCTCAGTAGCAGGACCAAGT 2160
DB |||||
DB 2101 GAAATTTGATCCCAACCCAGTGAAGCAGCTACTCTCTCAGTAGCAGGACCAAGT 2160
QY |||||
QY 2161 CCCTCGGGGGGAAAGTGTGTCATCAGTGTGAGTGGGGTTCCAGATTAATCCCTTAAACT 2220
DB |||||
DB 2161 CCCTCGGGGGGAAAGTGTGTCATCAGTGTGAGTGGGGTTCCAGATTAATCCCTTAAACT 2220
QY |||||
QY 2221 CATAGACCCAGTCTCCAGACAGAACAGTATTTGGAATATCCCGAGTGGCCAGTTCCT 2280
DB |||||
DB 2221 CATAGACCCAGTCTCCAGACAGAACAGTATTTGGAATATCCCGAGTGGCCAGTTCCT 2280
QY |||||
QY 2281 GCTTCTCCCAAGCTGGTATTATTCCAAACCTGAAGCTGTGGCAATGTCCTTAAACT 2340
DB |||||
DB 2281 GCTTCTCCCAAGCTGGTATTATTCCAAACCTGAAGCTGTGGCAATGTCCTTAAACT 2340
QY |||||
QY 2341 TCTCTCTAGCCCTGCTAGTCTCTGATCGAACAGAGCTGTATCCCTTCTAGTGGCTTAA 2400
DB |||||
DB 2341 TCTCTCTAGCCCTGCTAGTCTCTGATCGAACAGAGCTGTATCCCTTCTAGTGGCTTAA 2400
QY |||||
QY 2401 GATTCAGGCTTCAAGATCAGAGGAGCAACTCTCTCTGAGGGGAATAAAGAAAAATTTAAA 2460
DB |||||
DB 2401 GATTCAGGCTTCAAGATCAGAGGAGCAACTCTCTCTGAGGGGAATAAAGAAAAATTTAAA 2460
QY |||||
QY 2461 CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAAGTGTGT 2520
DB |||||
DB 2461 CCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAAGTGTGT 2520
QY |||||
QY 2521 TCTGAAACATAGAAACAGATGATGATTTTAAAGAAATTTAAGAAATGATTTTGGTTACAG 2580
DB |||||

DB 2521 TCTGAAACATAGAAACAGATGATGATTTTAAAGAAATTTAAGAAATGATTTTGGTTACAG 2580
QY |||||
QY 2581 CCAAGTTCTACTTCTGAAATCTATGATCAACTACTTAAACAAAAATAGAGAGGAGAAAA 2640
DB |||||
DB 2581 CCAAGTTCTACTTCTGAAATCTATGATCAACTACTTAAACAAAAATAGAGAGGAGAAAA 2640
QY |||||
QY 2641 TCAAGAGATTTGATCAAAGACAAAAATTTGAACAAAGTCTAAGGATTTCTTTCAATTTGAAAT 2700
DB |||||
DB 2641 TCAAGAGATTTGATCAAAGACAAAAATTTGAACAAAGTCTAAGGATTTCTTTCAATTTGAAAT 2700
QY |||||
QY 2701 AGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCGGAATAGCCCCAGCATTTCCCTTCA 2760
DB |||||
DB 2701 AGCAGCAGCAACTGTGTACAGTGGCAGCAGCAAGCGGAATAGCCCCAGCATTTCCCTTCA 2760
QY |||||
QY 2761 ATACTTTAGTAACCGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACT 2820
DB |||||
DB 2761 ATACTTTAGTAACCGGAGCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACT 2820
QY |||||
QY 2821 TCCAGCCAGCATGTAAACAAAGAGAAAGACGATAGGAAGAGAAAGAGACGAGCTGAG 2880
DB |||||
DB 2821 TCCAGCCAGCATGTAAACAAAGAGAAAGACGATAGGAAGAGAAAGAGACGAGCTGAG 2880
QY |||||
QY 2881 CAAGTTAGGAAATCAACATTTGAAATCCCAATGCAAGAGGAGTTCAACCCACGTTCTCTCT 2940
DB |||||
DB 2881 CAAGTTAGGAAATCAACATTTGAAATCCCAATGCAAGAGGAGTTCAACCCACGTTCTCTCTCT 2940
QY |||||
QY 2941 CAGCAAGAGCTTCTACTTACCCCAACTTTCAGCTCGGCTCAAGCAGCAACCTAGGCCATCT 3000
DB |||||
DB 2941 CAGCAAGAGCTTCTACTTACCCCAACTTTCAGCTCGGCTCAAGCAGCAACCTAGGCCATCT 3000
QY |||||
QY 3001 ATGGTGGTCTATCAACAGCCAACTCCAGTTTATATCTCAGCCCTGTTTGTGTTTGCACCAAT 3060
DB |||||
DB 3001 ATGGTGGTCTATCAACAGCCAACTCCAGTTTATATCTCAGCCCTGTTTGTGTTTGCACCAAT 3060
QY |||||
QY 3061 ATGATGATCCAGTCCAGTGGCCAGCCAGCGTGCAACTTTATATCTCAGCCCTGTTTGTGTTTGCACCAAT 3120
DB |||||
DB 3061 ATGATGATCCAGTCCAGTGGCCAGCCAGCGTGCAACTTTATATCTCAGCCCTGTTTGTGTTTGCACCAAT 3120
QY |||||
QY 3121 CCCATGCCAGTGAATCAAGCAGCAGACATATAGAGCAGTACCAGTATGCCCCAAGAGGG 3180
DB |||||
DB 3121 CCCATGCCAGTGAATCAAGCAGCAGACATATAGAGCAGTACCAGTATGCCCCAAGAGGG 3180
QY |||||
QY 3181 CAAGAACAAGCATATCAGAGTGCATGATGCAACCCAGCGTCAGCAGCGGGGCCACCGATT 3240
DB |||||
DB 3181 CAAGAACAAGCATATCAGAGTGCATGATGCAACCCAGCGTCAGCAGCGGGGCCACCGATT 3240
QY |||||
QY 3241 GCAGCCACCCACAGCTTACTCCAGGCAATATGTTGGCTACAGTCTCTCAGCAGTTCCTCA 3300
DB |||||
DB 3241 GCAGCCACCCACAGCTTACTCCAGGCAATATGTTGGCTACAGTCTCTCAGCAGTTCCTCA 3300
QY |||||
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
DB |||||
DB 3301 AATCAGCCCTTGTTCAGCATGTGCCATATATCAGTCTCAGCATCTCTCATGTCTATAGT 3360
QY |||||
QY 3361 CCTGTAATACAGGTAATGCTTAGAATGATGGCAACCAACACACAGCCCGCCTGGTTTA 3420
DB |||||
DB 3361 CCTGTAATACAGGTAATGCTTAGAATGATGGCAACCAACACACAGCCCGCCTGGTTTA 3420
QY |||||
QY 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGAGCGCATGCGATGATGCAATGT 3480
DB |||||
DB 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGAGCGCATGCGATGATGCAATGT 3480
QY |||||
QY 3481 CCCAAATTTACCATCAACAGGAGCAAGCCCTTCTTTCTACTTTGCAATTTCCACGGGC 3540
DB |||||
DB 3481 CCCAAATTTACCATCAACAGGAGCAAGCCCTTCTTTCTACTTTGCAATTTCCACGGGC 3540
QY |||||
QY 3541 TCCCTTGTCTCAGCAGTATGCGCACTTAAAGCTTACCTTGCACCCACATATCTTCCACCT 3600
DB |||||
DB 3541 TCCCTTGTCTCAGCAGTATGCGCACTTAAAGCTTACCTTGCACCCACATATCTTCCACCT 3600
QY |||||
QY 3601 CAGCTTTCAGCTACCCCACTGGGACAGGAGCAAGCCAAATGTTGGAGTTCATCTCTGCA 3660
DB |||||

3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCAACATGGTGGAGTCACTCTGCA 3660
3661 CCAGTCTCTGTTTCAGACCATCAGCACCAGGCGCGCCAGGCTCTCCATCTGCGCCAGTCCA 3720
3661 CCAGTCTCTGTTTCAGCACCATCAGCACCAGGCGCGCCAGGCTCTCCATCTGCGCCAGTCCA 3720
3721 CAGCAGCAGTCAAGCAATTTACCAAGCGGGGCTTGGCGCACTCCACCTCCATGACACT 3780
3721 CAGCAGCAGTCAAGCAATTTACCAAGCGGGGCTTGGCGCACTCCACCTCCATGACACT 3780
3781 GCCTCCAAACAGCAGTCCGACAGATAGTTTCCAGCAGCAGCAACAGACTGTCTTTAG 3840
3781 GCCTCCAAACAGCAGTCCGACAGATAGTTTCCAGCAGCAGCAACAGACTGTCTTTAG 3840
3841 ATCCATCTCTTCTCACTGTCAGCGGGGTATACCAACCCACCCACATGCGCCACGTA 3900
3901 CAGGCTCATGTACAGTCAGGAATGGTCTTCTCATCCAACTGCCCATGGCCAAATGATG 3960
3901 CAGGCTCATGTACAGTCAGGAATGGTCTTCTCATCCAACTGCCCATGGCCAAATGATG 3960
3961 CTAATCAGCAGACAGCAGCAGCGCGGTCCCGAGCGCGCTGCTCAAAAGTGCACTACAG 4020
3961 CTAATCAGCAGACAGCAGCAGCGCGGTCCCGAGCGCGCTGCTCAAAAGTGCACTACAG 4020
4021 CCATTTCCAGTCTCGACAGCGCATTTCCCTATATGACGACCCCTTCAGTACAAAGCC 4080
4021 CCATTTCCAGTCTCGACAGCGCATTTCCCTATATGAGGCAACCTTCAGTACAAAGCC 4080
4081 CACCAACACAGCAGTGTAAAGCTGCTGAGAGGAAACGAAAGGCAAAATTCCTCTCCTC 4140
4081 CACCAACACAGCAGTGTAAAGCTGCTGAGAGGAAACGAAAGGCAAAATTCCTCTCCTC 4140
4141 CTTTCTACTGCTTCTACCACTCGAAGCAGCAGAAACTAGAAATTCATTTATTTGTTTTT 4200
4141 CTTTCTACTGCTTCTACCACTCGAAGCAGCAGAAACTAGAAATTCATTTATTTGTTTTT 4200
4201 TAAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAATGTAACAGTTTCACTTGCAG 4260
4201 TAAATATATATGTTGATTTCTTTGTAACATCCAAATAGGAATGTAACAGTTTCACTTGCAG 4260
4261 TGAAGATATCTGACCGAGTAGAGGCAATTTAGGAACTTTGGGGGCTATTTCATTAATCCA 4320
4261 TGAAGATATCTGACCGAGTAGAGGCAATTTAGGAACTTTGGGGGCTATTTCATTAATCCA 4320
4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTGCGGAACTGGAGTTATTT 4380
4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTGCGGAACTGGAGTTATTT 4380
4381 ATTTTATTAATAACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAACAAGAGT 4440
4381 ATTTTATTAATAACCTTTGAAAGTCATGAACACATCAGCTAGCAAAAGAAAGTAACAAGAGT 4440
4441 GATTTGCTGCTATTTACTCTCTAAAAAATAAAAAAAAAAAAAA 4481
4441 GATTTGCTGCTATTTACTCTCTAAAAAATAAAAAAAAAAAAAA 4481

RESULT 6
US-09-949-016-4498
; Sequence 4498, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4498
; LENGTH: 4484
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4498

Query Match
Best Local Similarity 99.2%; Score 4447; DB 3; Length 4484;
Matches 4475; Conservative 0; Mismatches 5; Indels 5; Gaps 2;

QY 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCCCGCTCTCAAGTGTCCCTTCCCGGCCCCCGG 60
DB 1 ACCCCCGAGAAAGCAACCCAGCGCGCGCGCCCGCTCTCAAGTGTCCCTTCCCGGCCCCCGG 60
QY 61 GCACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
DB 61 GCACCTCAGCTTCTGCTTCCGTCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
QY 121 CACCTCCGCTCCACCCCGCGGCTCGGGCGCGCCCTCCGATCGCTCAGCGGCGCA 180
DB 121 CACCTCCGCTCCACCCCGCGGCTCGGGCGCGCCCTCCGATCGCTCAGCGGCGCA 180
QY 181 GCTCTCCGAGTCCCGCGGTGGCCACCGAGTCTCGCGTTCGCCGACGCCAGGTGGCCC 240
DB 181 GCTCTCCGAGTCCCGCGGTGGCCACCGAGTCTCGCGTTCGCCGACGCCAGGTGGCCC 239
QY 241 GGGTGGCGTCTGCTCAGCGCGCGCGCGCGGCGGAGCGGGCGGGCGGGTGGCGGGCC 300
DB 241 GGGTGGCGTCTGCTCAGCGCGCGCGCGCGGCGGAGCGGGCGGGCGGGTGGCGGGCC 299
QY 301 CCGGAGCGTATCTCTCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 CCGGAGCGTATCTCTCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359
QY 361 CAGAGCTCGCTCTCCCTCCGCTCAGACTGTGTTTGTAGCAAGCGCAACCGCGCGCGCG 420
DB 361 CAGAGCTCGCTCTCCCTCCGCTCAGACTGTGTTTGTAGCAACCGCGCGCGCGCGCG 419
QY 421 TTTTGGCGCGCTCGCGCGCGCTCTGCTGCTGCGGGCGCTTCCCGCGCGCGCTTCTGTCGTC 480
DB 421 TTTTGGCGCGCTCGCGCGCGCTCTGCTGCTGCGGGCGCTTCCCGCGCGCGCTTCTGTCGTC 479
QY 481 GTCTTCTCTCCCTCGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 GTCTTCTCTCCCTCGCGCGCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCG 539
QY 541 CTCGGCGCGCGTGGTCCCGCGCGCTTCCGGCGCTCTCTTGGCGCGCGCGCGCGCGCG 600
DB 541 CTCGGCGCGCGTGGTCCCGCGCGCTTCCGGCGCTCTCTTGGCGCGCGCGCGCGCGCG 599
QY 601 TGTTCCTCCCGCGCGTGGTGGCGCGCTTATGGCGCGCTTCAACATGTCTGCTGAAGCCCCAG 660
DB 601 TGTTCCTCCCGCGCGTGGTGGCGCGCTTATGGCGCGCTTCAACATGTCTGCTGAAGCCCCAG 659
QY 661 CAG 720
DB 661 CAG 719
QY 721 CAGCGCGCGCGCGCGCTGCTCAATGTCCGAAAGCGCGCGCGCGCGCGCGCTTCTTAGCGTCG 780
DB 721 CAGCGCGCGCGCGCGCTGCTCAATGTCCGAAAGCGCGCGCGCGCGCGCTTCTTAGCGTCG 779
QY 781 CCGCGCGCGCGCGCTTCCCGCGCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCC 840
DB 781 CCGCGCGCGCGCGCTTCCCGCGCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCC 839
QY 841 TCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900

Dd	840	TCCTCGTGTGTCGGCGGACCTTCGGCGGGAGGCCCGGCTTGGGCAGAGGTGGAAC	899	Qy	1977	TCCTGTCTTACTATGCTTAAACGCATGTCTTCAAGAGGGCTCCAAAGGATGTCCCCAA	2036
Qy	901	AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGATGGAATCTATGCAAAATGAGG	960	Dd	1980	TCCTGTCTTACTATGCTTAAACGCATGTCTTCAAGAGGGCTCCAAAGGATGTCCCCAA	2039
Dd	900	AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGATGGAATCTATGCAAAATGAGG	959	Qy	2037	GGCCAGCGCATCTCGAATACACAGATTTCTGCTGGGAGGGTTCATATCCAGTGG	2096
Qy	961	ATGCTTCATATCTTACATCAGTCTTGCTCCAAATGTGAATGTAAGTGAATAATGGA	1020	Dd	2040	GGCCAGCGCATCTCGAATACACAGATTTCTGCTGGGAGGGTTCATATCCAGTGG	2099
Dd	960	ATGGTTCATATCTTACATCAGTCTTGCTCCAAATGTGAATGTAAGTGAATAATGGA	1019	Qy	2097	CCTAGAAATTTGTATCCCAACACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGAC	2156
Qy	1021	GGTATATGAAGGAGTTTAAACTTACAGTCCGAAGTGTGATTTGCTACTTCATGCC	1080	Dd	2100	CCTAGAAATTTGTATCCCAACACCCAGTGAAGCAGCTACTCTCCAGTAGCAAGAC	2159
Dd	1020	GGTATATGAAGGAGTTTAAACTTACAGTCCGAAGTGTGATTTGCTACTTCATGCC	1079	Qy	2157	CAGTCCCTCGGGGGAACGTGGTTCATCAGTGGTTCAGTGGGTTCGAAGATTTATCCCTAA	2216
Qy	1081	GCACATGAGAAAGTACAGAATCCAGTTCGGGGCCGAACCGTGAAGAAATATGAGAGT	1140	Dd	2160	CAGTCCCTCGGGGGAACGTGGTTCATCAGTGGTTCAGTGGGTTCGAAGATTTATCCCTAA	2219
Dd	1080	GCACATGAGAAAGTACAGAATCCAGTTCGGGGCCGAACCGTGAAGAAATATGAGAGT	1139	Qy	2217	AACTCATAGACCCAGGTCTCCACAGACAGACAGTATTTGGAATACCCCGAGTGGGCCAGT	2276
Qy	1141	ATTTTGTTCAAATGTTACAGCTTTGTTGTGTGACAGTTTAAAGATATGACTCCAGTTAT	1200	Dd	2220	AACTCATAGACCCAGGTCTCCACAGACAGACAGTATTTGGAATACCCCGAGTGGGCCAGT	2279
Dd	1140	ATTTTGTTCAAATGTTACAGCTTTGTTGTGTGACAGTTTAAAGATATGACTCCAGTTAT	1199	Qy	2277	TCTTGTCTTCTCCCAAGCTGTTATTTCCAACTGAAAGCTGTTGCCATGCTATTTCAGC	2336
Qy	1201	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTGCTTAAAGTGAATGGCGAACACAA	1260	Dd	2280	TCTTGTCTTCTCCCAAGCTGTTATTTCCAACTGAAAGCTGTTGCCATGCTATTTCAGC	2339
Dd	1200	GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTGCTTAAAGTGAATGGCGAACACAA	1259	Qy	2337	TGCATCTCTACGCTGCTAGTCTCTGATCGAACAGAGCTGTTTACCCCTTCTAGTGAGGC	2396
Qy	1261	GAGAAAGCCTGGAGCCCTGGGATCGAGTGAATCTCACAGCAATGAGGAACTTGAGGCT	1320	Dd	2340	TGCATCTCTACGCTGCTAGTCTCTGATCGAACAGAGCTGTTTACCCCTTCTAGTGAGGC	2399
Dd	1260	GAGAAAGCCTGGAGCCCTGGGATCGAGTGAATCTCACAGCAATGAGGAACTTGAGGCT	1319	Qy	2397	TAAAGATTTCCAGGCTTCAAGATCAGAGCGAGAACTCTCTGCGAGGGAATAAGAAAAATAT	2456
Qy	1321	TTGGAAATGACGTATCTAATGATGGATCCCAATGATATGTTTCGATATATAAGAA	1380	Dd	2400	TAAAGATTTCCAGGCTTCAAGATCAGAGCGAGAACTCTCTGCGAGGGAATAAGAAAAATAT	2459
Dd	1320	TTGGAAATGACGTATCTAATGATGGATCCCAATGATATGTTTCGATATATAAGAA	1379	Qy	2457	TAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACCAAGGTATATCACCAGT	2516
Qy	1381	AATTATGTGTAGTGTACGTATGATAGCAGTTTATCTTGTATACAGTGCCTTAGAA	1440	Dd	2460	TAAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACCAAGGTATATCACCAGT	2519
Dd	1380	AATTATGTGTAGTGTACGTATGATAGCAGTTTATCTTGTATACAGTGCCTTAGAA	1439	Qy	2517	TGTTTCTGAAATAGAAAAACAGATTTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGGTT	2576
Qy	1441	AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAA	1500	Dd	2520	TGTTTCTGAAATAGAAAAACAGATTTGATGATTTTAAAGAAATTTTAAAGATGATTTTAGGTT	2579
Dd	1440	AGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGGCAAAACAGTTAGCAGAA	1499	Qy	2577	ACGCCAGTTCTACTTCTGATCTATGGATCAACTACTTAACAAAAATAGAGAGGAGA	2636
Qy	1501	ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAATGATGATAGGAGTAG	1560	Dd	2580	ACGCCAGTTCTACTTCTGATCTATGGATCAACTACTTAACAAAAATAGAGAGGAGA	2639
Dd	1500	ATTGAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAATGATGATAGGAGTAG	1559	Qy	2637	AAAAATCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTCTTAAGGATTTCTTATTGA	2696
Qy	1561	GAAGAAAAATACACAGCAGTTCCAGAGAAATCCAGTGAACGTGAGGGGCACAGCATAAC	1620	Dd	2640	AAAAATCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTCTTAAGGATTTCTTATTGA	2699
Dd	1560	GAAGAAAAATACACAGCAGTTCCAGAGAAATTTCCAGTGAACGTGAGGGGCACAGCATAAC	1619	Qy	2697	AAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCC	2756
Qy	1621	ACTAGGAAAAATATAATATTTCTCTGGACAAAGAAATAGAGAAAGTCAATATCTTGGGGA	1680	Dd	2700	AAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTCCCC	2759
Dd	1620	ACTAGGAAAAATATAATATTTCTCTGGACAAAGAAATAGAGAAAGTCAATATCTTGGGGA	1679	Qy	2757	TTCAATCTTAGTTAAACAGGAGCAACAGGGGACCTGAGGTCACTTCCCAAGGGGTTCA	2816
Qy	1681	AGTGGAGACAGAAATCACCGGTATGGCCAGCCTGGATGGGCTCCATGCCATCAAGA	1740	Dd	2760	TTCAATCTTAGTTAAACAGGAGCAACAGGGGACCTGAGGTCACTTCCCAAGGGGTTCA	2819
Dd	1680	AGTGGAGACAGAAATCACCGGTATGGCCAGCCTGGATGGGCTCCATGCCATCAAGA	1739	Qy	2817	GACTTCAGCCCGCAGCATGTAAACAGAGAAAGACGATAAGGAAGAGAAAGAACGACG	2876
Qy	1741	TCCACTTCTCACACTTCAGATTTCAACCCGAATTTCTGGTTTCAGACCAAGAGTAGTTAAT	1800	Dd	2820	GACTTCAGCCCGCAGCATGTAAACAGAGAAAGACGATAAGGAAGAGAAAGAACGACG	2879
Dd	1740	TCCACTTCTCACACTTCAGATTTCAACCCGAATTTCTGGTTTCAGACCAAGAGTAGTTAAT	1799	Qy	2877	TGACCAAGTTAGGAAATCAACATTTGAATTCGCAATGCAAAAGAGGTTCAACCCACGTTCTTT	2936
Qy	1801	GGAG----GTGTTCCCTGGCCATCGCCTTGCCCATCTCTCTCTCTCGCCCACTTCTCG	1856	Dd	2880	TGACCAAGTTAGGAAATCAACATTTGAATTCGCAATGCAAAAGAGGTTCAACCCACGTTCTTT	2939
Dd	1800	GGAGCAAGTGTCTCTGGCCATCGCCTTGCCCATCTCTCTCTCTCGCCCACTTCTCG	1859	Qy	2937	CTCTCAGCCAAAGCTTCTACTACCCCACTTCACTTCCGCTCGGCTCAAGCACAACCTAGCCC	2996
Qy	1857	CTACAGTCAGTCCCAACTCTCTCCAGCTCGGCAGCACCCTTACACGCGCCCTC	1916	Dd	2940	CTCTCAGCCAAAGCTTCTACTACCCCACTTCACTTCCGCTCGGCTCAAGCACAACCTAGCCC	2999
Dd	1860	CTACAGTCAGTCCCAACTCTCTCCAGCTCGGCAGCACCCTTACACGCGCCCTC	1919	Qy	2997	ATCTATGTTGGGTTCATCAACAGCCAACTCCAGTTTATATCTCAGGCTGTTTGTGTTGAC	3056
Qy	1917	CAGGCCCTCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGTCATGTTTCTCCAGC	1976	Dd	3000	ATCTATGTTGGGTTCATCAACAGCCAACTCCAGTTTATATCTCAGGCTGTTTGTGTTGAC	3059
Dd	1920	CAGGCCCTCTCGGGCCATCCAGACCCCGTCTCACCCCTCTGTCATGTTTCTCCAGC	1979				

QY 3057 AAATATGATGATCCAGTCCAGTCCAGCCAGCGGTGCAACCTTTTATACCAATACCTAT 3116
DB 3060 AAATATGATGATCCAGTCCAGTCCAGCCAGCGGTGCAACCTTTTATACCAATACCTAT 3119
QY 3117 GACGCCATGCGCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCAAAATATATGCCCAACA 3176
DB 3120 GACGCCATGCGCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCAAAATATATGCCCAACA 3179
QY 3177 GCGGCAAGACAGCAGTATCATGAGTGCATGATGACCCAGCGTCAAGCGGGGCCAC 3236
DB 3180 GCGGCAAGACAGCAGTATCATGAGTGCATGATGACCCAGCGTCAAGCGGGGCCAC 3239
QY 3237 GATTGAGCAGCCACCAAGCTTACTCCAGCGAATATGTTGCTCAGTCCCTCAGCAGTT 3296
DB 3240 GATTGAGCAGCCACCAAGCTTACTCCAGCGAATATGTTGCTCAGTCCCTCAGCAGTT 3299
QY 3297 CCCAATCAGCCCTTGTTCAGCATGTGCCATATATCATGTCTCAGCATCTCTCATGTCTTA 3356
DB 3300 CCCAATCAGCCCTTGTTCAGCATGTGCCATATATCATGTCTCAGCATCTCTCATGTCTTA 3359
QY 3357 TAGTCTGTAATACAGGGTAATGCTAGAAATGATGCGACCAACACAGCGCCAGCTGG 3416
DB 3360 TAGTCTGTAATACAGGGTAATGCTAGAAATGATGCGACCAACACAGCGCCAGCTGG 3419
QY 3417 TTTAGTATCTTCTTACGCACTCAGTACGGGGCTCATGAGCAGCGCATGCGATGTATGC 3476
DB 3420 TTTAGTATCTTCTTACGCACTCAGTACGGGGCTCATGAGCAGCGCATGCGATGTATGC 3479
QY 3477 ATGTCCCAATATACCATACAAAGGAGACAGCCCTTCTTCTACTTTGGCCATTTCCAC 3536
DB 3480 ATGTCCCAATATACCATACAAAGGAGACAGCCCTTCTTCTACTTTGGCCATTTCCAC 3539
QY 3537 GGGCTCCCTTGTCTCAGCAGTATGCGACCCCTAACCGTACCTGACCCACATCTCCACA 3596
DB 3540 GGGCTCCCTTGTCTCAGCAGTATGCGACCCCTAACCGTACCTGACCCACATCTCCACA 3599
QY 3597 CCTCAGCCTTACGTTACCCCACTGGACAGCAGCAAGCAACATGTTGGAGTCAATCC 3656
DB 3600 CCTCAGCCTTACGTTACCCCACTGGACAGCAGCAAGCAACATGTTGGAGTCAATCC 3659
QY 3657 TGCACCCAGTCTGTTACGACCATCAGCAGCAGCGCGCCAGCGTCTCATCTGCGCCAG 3716
DB 3660 TGCACCCAGTCTGTTACGACCATCAGCAGCAGCGCGCCAGCGTCTCATCTGCGCCAG 3719
QY 3717 TCCACAGCAGCAGTACGCAATTTACACGCGGGCTTGGCCCACTCCACCTCCATGAC 3776
DB 3720 TCCACAGCAGCAGTACGCAATTTACACGCGGGCTTGGCCCACTCCACCTCCATGAC 3779
QY 3777 ACCTGCTTCAACACGAGTTCGCGCACAGAAATAGTTTCCAGCAGCAGCAACAGCTGTCTT 3836
DB 3780 ACCTGCTTCAACACGAGTTCGCGCACAGAAATAGTTTCCAGCAGCAGCAACAGCTGTCTT 3839
QY 3837 TAGATCCATCTTCTCAGCTTCCAGCGGGGTATACCAACCCACCCACCATGCGCCACCT 3896
DB 3840 TAGATCCATCTTCTCAGCTTCCAGCGGGGTATACCAACCCACCCACCATGCGCCACCT 3899
QY 3897 ACCTCAGGCTCATGTACAGTACAGTATGTTTCTTCTCCTCAACTGCGCATGCGCAAT 3956
DB 3900 ACCTCAGGCTCATGTACAGTACAGTATGTTTCTTCTCCTCAACTGCGCATGCGCAAT 3959
QY 3957 GATGCTAATGACGACACAGCCACCGCGGTTCGCCAGCGCGCTCGCTCAAGTGCAT 4016
DB 3960 GATGCTAATGACGACACAGCCACCGCGGTTCGCCAGCGCGCTCGCTCAAGTGCAT 4019
QY 4017 ACAGCCATTCAGTCTCAGCAACAGGCGATTTCCCTATATACCGCACCTTCAGTACA 4076
DB 4020 ACAGCCATTCAGTCTCAGCAACAGGCGATTTCCCTATATACCGCACCTTCAGTACA 4079
QY 4077 AGCCACACCAACAGCAGTTGTAAGCTCCTCGGAGGAACGGAAGGCGCAATTCCT 4136
DB 4080 AGCCACACCAACAGCAGTTGTAAGCTCCTCGGAGGAACGGAAGGCGCAATTCCT 4139
QY 4137 CCTCCCTTCTACTGCTTCTACCACTGGAAAGCAGAGAAACTAGAAATTTATTTTG 4196

DB 4140 CCTCCCTTCTACTGCTTCTACCACTGGAGCAGCAAGAACTAGAAATTTATTTTG 4199
QY 4197 TTTTAAATATATATGTTGATTTCTTGTAATCATCAATAGGAATGCTAAGTTCACCT 4256
DB 4200 TTTTAAATATATATGTTGATTTCTTGTAATCATCAATAGGAATGCTAAGTTCACCT 4259
QY 4257 GCAGTGGAGATACCTGACCGAGTAGAGGCAATTTAGGAACCTTGGGGCTATTCATAT 4316
DB 4260 GCAGTGGAGATACCTTGGACGAGTAGAGGCAATTTAGGAACCTTGGGGCTATTCATAT 4319
QY 4317 TCCATATCTGCTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTCCGAAACTGGAAGTT 4376
DB 4320 TCCATATCTGCTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTCCGAAACTGGAAGTT 4379
QY 4377 ATTTATTTTAAATAACCCCTTGAAGTCAATGAACACATCAGCTAGCAAAAGTAAACAA 4436
DB 4380 ATTTATTTTAAATAACCCCTTGAAGTCAATGAACACATCAGCTAGCAAAAGTAAACAA 4439
QY 4437 GAGTGAATTTGCTGCTATTACTGCTAAATAAAAAAAAAAAAAA 4481
DB 4440 GAGTGAATTTGCTGCTATTACTGCTAAATAAAAAAAAAAAAAA 4484

RESULT 7
US-08-981-998A-4
; Sequence 4, Application US/08981998A
; Patent No. 6844431
; GENERAL INFORMATION:
; APPLICANT: PULST, STEFAN M.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCERECELLAR
; ;
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUEITING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: 55401
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,998A
; FILING DATE: 11-May-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 97/42314
; FILING DATE: 08-MAY-1997
; APPLICATION NUMBER: US 08/727,084
; FILING DATE: 08-OCT-1996
; APPLICATION NUMBER: US 60/022,207
; FILING DATE: 19-JUL-1996
; APPLICATION NUMBER: US 60/017,388
; FILING DATE: 08-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MUEITING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 232.00010120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3798 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:

[illegible]

```

RESULT 8
US-09-648-281-11
; Sequence 11, Application US/09648281
; Patent No. 6515197
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M.
; TITLE OF INVENTION: Transgenic Animal Model of
; TITLE OF INVENTION: Neurodegenerative Disease and Methods of Use
; FILE REFERENCE: P-CE 4336
; CURRENT APPLICATION NUMBER: US/09/648,281
; CURRENT FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2) .. (1257)
US-09-648-281-11

```

	Query Match	20.9%;	Score	935.2;	DB 3;	Length	1257;	
	Best Local Similarity	85.9%;	Pred. No.	6e-195;				
	Matches 1062; Conservative	0;	Mismatches	168;	Indels	6;	Gaps	2;
QY	685	CAGCAGCAGCAGCAACGACGACGACGACGACGACGACGCGCCGCCGGTGTGCCAAT	744					
DB	2	CACGAGGGGCGCTTCACCATGTCTCTGAAGCGGAGCGAGCGCGCGCGCGCGGCAC	61					


```

Db 242 CCTACGATTTCTTTGATGGAATCTATGCAACGTGAGATGGTTCATATATCTTACGTCA 301
Qy 982 GTTGTGGCTCAAAATGTGAAGTACAAGTGAATAATGAGAGTATATATCAAGAGTGTCTT 1041
Db 302 GTTGTGGATCGAAATGTGAAGTACAAGTGAATAATGAGAGTATATATGAAGAGTGTCTT 361
Qy 1042 AAACTTACAGTCCGAAAGTGTGATTTGGTATCTTGATGCGGCACATGAGAAAAGTACAGAA 1101
Db 362 AAAACATACAGTCCTAAGTGTGACTTGGTACTTGTGCTGCACATGAGAAAAGTACAGAA 421
Qy 1102 TCCAGTTCGGGGCCGAAAGCTGAAGAAATAATGAGAGTATTTCTCAATGTTCAGAC 1161
Db 422 TCCAGTTCGGGGCCGAAAGCTGAAGAAATAATGAGAGTGTCTTGTCAATGTCTCAGAC 481
Qy 1162 TTTGTGTGGTACAGTTAAAGATATGGAATCCAGTTCAGTATGCAAAAAGAGATGCTTTTACT 1221
Db 482 TTCGTTGTGGTACAGTTTAAAGATACAGACTCCAGTTATGACGGAGAGATGCTTTTACT 541
Qy 1222 GACTCTGCTATCAGTGTCTTAAAGTGAATGCGGAACAACAAGAGAGGACCTGGAGCCCTGG 1281
Db 542 GACTCTGCTCTCAGCGCAAGAGTGAATGCTGAGCAACAAGGAGAGGACCTGGAGCCCTGG 601
Qy 1282 GATGCAAGTGAATCACAGCCCAATGAGGAATCTGAGGCTTTGGAAATGACGTATCTAAT 1341
Db 602 GATGCAAGGAGACTCACGGCCAGCAGGAGCTGGAG---CTGGAGAAATGATGTGTCTAAT 658
Qy 1342 GGATGGGATCCCAATGATATGTTTCGATATATGAAGAAAATATGCTGTAGTGTCTACG 1401
Db 659 GGATGGGACCCCAATGACATGTTTCGATATATGAAGAAATATGCTGTAGTGTCTCACA 718
Qy 1402 TATGATAGCAGTTTATCTTCGATATACAGTGCCTTTAGAAAAGAGATAACTCAGAGAAATTT 1461
Db 719 TATGATAGCAGTTTATCTTCATATACGCTTCCTTTAGAAAAGGACAACTCAGAGAAATTT 778
Qy 1462 TTAAACCGGAGAGCAAGGCAACACAGTTAGCAGAGAAATTTGATCAAGTCCCGAGTAC 1521
Db 779 CTAAACCGGAGGCAAGGCAACACAGTTAGCAGAGAAATTTGAATCCAGTGTCTCAGTAC 838
Qy 1522 AAAGCTCGAGTGGCCCTTGAAAATGATGATGAGAGTGAAGAAAATAACACAGCAAT 1581
Db 839 AAAGCTCGTGTGCGCTTTGAGATGATGACCGGAGTGAAGAAAATAACACAGCAATC 898
Qy 1582 CAGAGAAATTCAGTGAACGTGAGGGGCACAGCATAAACACTAGGAAAATAAATATATTT 1641
Db 899 CAGAGAAATCTCAGTGAACCGGAGGGGCATGCCCCCAACACTAGGCACAATAAATATATTT 958
Qy 1642 CCTCTTGACAAAGAAATAGAGAACTCATCTCGGGAAGTGGGAGACAGAAATTCACCG 1701
Db 959 CCTCTTGACAAAGAAACAGAGAAATGCTCATCTCGGGAAGTGGGAGACAGAGCTCACCA 1018
Qy 1702 CGTATGGGCGAGCTGGATCGGGCTCCATGCGCATCAAGATCCACTTCTCACACTTCAGAT 1761
Db 1019 CGGATGGGCGAGCTGGGCGAGGCTCCATGCGCTCAGAGTGTCTTCTCACACTTCAGAT 1078
Qy 1762 TTCAACCCGAAATTCGTTGTTCAGACCAAGAGTAGTTAATGAGAGTGTTCCTTGGCCATCG 1821
Db 1079 TTCAACCCGAAACGCTGGGCTCAGACCAAGAGTAGTTAATGAGAGTGTTCCTTGGCCATCG 1138
Qy 1822 CCTTGGCCATCTCTCTCTCGCCCACTTCTCGCTACCACTCAGTCCCACTCTT 1881
Db 1139 CCTTGGCCATCTCTCTCTCGCCCACTTCTCGCTACCACTCAGTCCCACTCTT 1198
Qy 1882 CCACCTCGGGCAGGCCACCCCTACAGGGCGCCCTCC 1917
Db 1199 CCACCTCGGGCAGGCCACCCCTACAGGGCTGTGTC 1234
```

RESULT 11
US-09-949-016-16240
; Sequence 16240, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16240
; LENGTH: 151088
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(151088)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16240
```

```

Query Match 19.4%; Score 870; DB 3; Length 151088;
Best Local Similarity 99.3%; Pred. No. 1.1e-179;
Matches 884; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 1 ACCCCCGAGAAAGCAACCCAGCGCGCCGCTCTCAGTGTCCTCCCGGCCCGGG 60
Db 2001 ACCCCCGAGAAAGCAACCCAGCGCGCCGCTCTCAGTGTCCTCCCGGCCCGGG 2060

Qy 61 GCCACCTCACGTTCTGCTTCCGTCGTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120
Db 2061 GCCACCTCACGTTCTGCTTCCGTCGTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 2120

Qy 121 CACTCTCGCTCCCAACCGCGGCCCTCGCGCGCCGCCCTCCGATGCGTCAAGCGGCCGA 180
Db 2121 CACTCTCGCTCCCAACCGCGGCCCTCGCGCGCCGCCCTCCGATGCGTCAAGCGGCCGA 2180

Qy 181 GCTCTCGGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTCGCGGACGAGCGGTGGCC 240
Db 2181 GCTCTCGGAGTCCCGCGGTGGCCACCGAGTCTCGCGG- TTCGCGGACGAGCGGTGGCC 2239

Qy 241 GGGTGGCGCTCGCTCCAGCGGCCGCGCGGAGCGGCGGGCGGGCGGGCGGCC 300
Db 2240 GGGTGGCGCTCGCTCCAGCGGCCGCGCGGAGCGGCGGGCGGGCGGGCGGGCGGCC 2299

Qy 301 CCGGAGCGTATCCCTCGCGCGCCCTCCCGCGCCCGGCCCGGCCCTTCCCTCCCGG 360
Db 2300 CCGGAGCGTATCCCTCGCGCGCCCTTCTCCGCGCCCGGCCCGGCCCTTCCCTCCCGG 2359

Qy 361 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGCAACGGCGGGCGGCG 420
Db 2360 CAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACGCAACGGCGGGCGGCGG 2419

Qy 421 TTTCCGCGCGGCTCCCGCGGCTCTTTGGTCTCGCGCGGCTCCCGGCCCTTCTGTCGTC 480
Db 2420 TTTCCGCGCGGCTCCCGCGGCTCTTTGGTCTCGCGCGGCTCCCGGCCCTTCTGTCGTC 2479

Qy 481 GTCCTTCTCCCTTCGCGAGCCCGGGCGCCCTTCGCGCGGCCCAACCGCGGCCCTCCCGG 540
Db 2480 CTCCTTCTCCCTTCGCGAGCCCGGGCGCCCTTCGCGCGGCCCAACCGCGGCCCTCCCGG 2539

Qy 541 CTCGCGCGCGTGGTCCCGCGCGGTTCGCGCGTCTTGGCGCGCCCGGCTCCCGG 600
Db 2540 CTCGCGCGCGCGTCCCGCGCGGTTCGCGCGTCTTGGCGCGCCCGGCTCCCGG 2599

Qy 601 TGTCCCGCGCGGTGCGAGCCGCTGTATGGGCGCCCTCACCATGTGCTGAAGCCCCAG 660
Db 2600 TGTCCCGCGCGGTGCGAGCCGCTGTATGGGCGCCCTCACCATGTGCTGAAGCCCCAG 2659

Qy 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720
```


Db 121 CTCGGCGCGCAACCGCGCTCCCGCTCGGCGCGCGTGTCTCCCGCGCGTTCGG 180
Qy 572 GCGTCTCTTGCGCGCGCGCGCTCCCGCTGTCTCCCGCGCGCGTGTCTCCCGCGCGTGTATG 631
Db 181 GCGTCTCTTGCGCGCGCGCGCTCCCGCTGTCTCCCGCGCGCGTGTCTCCCGCGCGTGTATG 240
Qy 632 GCGCGCTCACCATGTCTGAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 691
Db 241 GCGCGCTCACCATGTCTGAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy 692 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 751
Db 301 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 360
Qy 752 AGCCCGCGCGCGCGCGCTTCTAGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 811
Db 361 AGCCCGCGCGCGCGCGCTTCTAGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Qy 812 CGGTCTCTCTGCTCG 871
Db 421 CGGTCTCTCTGCTCG 480
Qy 872 GGAGCG 890
Db 481 GGAGCG 499

RESULT 14

US-08-981-998A-1
; Sequence 1, Application US/08981998A
; Patent No. 684431
; GENERAL INFORMATION:
; APPLICANT: PULST, STEFAN M.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
; ATAXIA-2 AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: 55401
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/981,998A
; FILING DATE: 11-May-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 97/42314
; FILING DATE: 08-MAY-1997
; APPLICATION NUMBER: US 08/727,084
; FILING DATE: 08-OCT-1996
; APPLICATION NUMBER: US 60/022,207
; FILING DATE: 19-JUL-1996
; APPLICATION NUMBER: US 60/017,388
; FILING DATE: 08-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MUETING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 232.00010120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 516 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both

; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-981-998A-1
Query Match 11.1%; Score 497.4; DB 3; Length 516;
Best Local Similarity 99.8%; Pred. No. 3.4e-99; Indels 0; Gaps 0;
Matches 498; Conservative 0; Mismatches 1;
Qy 392 TTGTAGCAACG 451
Db 1 TTGTAGCAACG 60
Qy 452 TCG 511
Db 61 TCG 120
Qy 512 CTCGGCG 571
Db 121 CTCGGCG 180
Qy 572 GCGTCTCTTGCG 631
Db 181 GCGTCTCTTGCG 240
Qy 632 GGCG 691
Db 241 GGCG 300
Qy 592 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 751
Db 301 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 360
Qy 752 AGCCCGCGCGCGCGCGCTTCTAGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 811
Db 361 AGCCCGCGCGCGCGCGCTTCTAGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Qy 812 CGGTCTCTCTGCTCG 871
Db 421 CGGTCTCTCTGCTCG 480
Qy 872 GGAGCG 890
Db 481 GGAGCG 499

RESULT 15

US-09-707-919A-13
; Sequence 13, Application US/09707919A
; Patent No. 6623927
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: Method for detection of human spinocerebeller ataxia 2
; TITLE OF INVENTION: gene variants
; FILE REFERENCE: US 443
; CURRENT APPLICATION NUMBER: US/09/707,919A
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Description of
; OTHER INFORMATION: Artificial derived nucleotide sequence of allelic
; OTHER INFORMATION: variant of SCA2 gene
US-09-707-919A-13

Query Match 10.2%; Score 455.8; DB 3; Length 459;
Best Local Similarity 99.6%; Pred. No. 4.2e-90;
Matches 457; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 375 CTCCGGCTCAGACTGTTTGTAGCAACGGCAACGGGGGGGGTTCGGCCCGGCTC 434
Db 1 CTCCGGCTCAGACTGTTTGTAGCAACGGCAACGGGGGGGGTTCGGCCCGGCTC 60
QY 435 CCGGCGGCTCCTTGGTCTCGGGGGGCTCCCGGCCCTTTCGTCTGTCTTCTTCTCCCT 494
Db 61 CCGGCGGCTCCTTGGTCTCGGGGGGCTCCCGGCCCTTTCGTCTGTCTTCTTCTCCCT 120
QY 495 CGCCAGCCCGGGGGCCCTTCGGCCCGGCCAAACCGCGCTCCCGCTCGGGCCCGGTC 554
Db 121 CGCCAGCCCGGGGGCCCTTCGGCCCGGCCAAACCGCGCTCCCGCTCGGGCCCGGTC 180
QY 555 GTCCCCCGCGGCTTCGGGGCTTCCTTGGGGGGCCCGGCTCCGGGCTGTCCCGCCCGGC 614
Db 181 GTCCCCCGCGGCTTCGGGGCTTCCTTGGGGGGCCCGGCTCCGGGCTGTCCCGCCCGGC 240
QY 615 GTCCGAGCCGGTGTATGGGCCCCCTCACCATGTCTGTAAGCCCGAGCAGCAGCA 674
Db 241 GTCCGAGCCGGTGTATGGGCCCCCTCACCATGTCTGTAAGCCCGAGCAGCAGCA 300
QY 675 GCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 734
Db 301 GCAGCAGCAACAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 360
QY 735 GGCTGCCAATGTCGGCAAGCCCGGGCAGCGGCTTCTAGGTCGCGCCCGCGCGGCGC 794
Db 361 GGCTGCCAATGTCGGCAAGCCCGGGCAGCGGCTTCTAGGTCGCGCCCGCGCGGCGC 420
QY 795 TTCGCGTCTCTCTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 833
Db 421 TTCGCGTCTCTCTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 459
```

Search completed: December 23, 2005, 05:15:14
Job time : 784 secs

This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	4481	100.0	4481	6	US-10-141-541-5		Sequence 5, Appli
2	4481	100.0	4481	8	US-10-150-323-2		Sequence 2, Appli
3	4481	100.0	4481	9	US-10-887-553A-493		Sequence 493, App
4	4481	100.0	4481	9	US-10-802-228-1		Sequence 1, Appli
5	3138.2	70.0	4481	6	US-10-141-541-3		Sequence 3, Appli
6	1730.6	38.6	2006	3	US-09-925-298-15		Sequence 15, Appl
7	1730.6	38.6	2006	5	US-10-102-806-15		Sequence 15, Appl
8	935.2	20.9	1257	6	US-10-141-541-1		Sequence 1, Appli
9	935.2	20.9	1257	8	US-10-150-323-4		Sequence 4, Appli
10	497.4	11.1	516	8	US-10-750-323-1		Sequence 1, Appli
11	452	10.1	682	9	US-10-956-157-3976		Sequence 3976, App
12	452	10.1	682	9	US-10-956-157-9211		Sequence 9211, App
13	441.2	9.8	704	8	US-10-363-345A-23185		Sequence 23185, A
14	441.2	9.8	704	8	US-10-363-345A-23186		Sequence 23186, A
15	441.2	9.8	704	9	US-10-363-483A-23185		Sequence 23185, A
16	441.2	9.8	704	9	US-10-363-483A-23186		Sequence 23186, A
17	437.6	9.8	485	3	US-09-918-995-28268		Sequence 28268, A
18	412.4	9.2	73995	5	US-10-087-192-208		Sequence 208, App
19	381	8.5	465	4	US-09-925-065A-129085		Sequence 129085,
20	301.4	6.7	305	7	US-10-242-335A-34580		Sequence 34580, A
21	301.4	6.7	305	7	US-10-085-783A-34580		Sequence 34580, A
22	288.2	6.4	493631	5	US-10-087-192-205		Sequence 205, App
23	281.4	6.3	285	4	US-09-925-065A-129084		Sequence 129084,

QY 301 CCGGACCGTATTCCTTCGCGCCGCCCTCCCCCGCCGCGCCCGGCCCTCCCTCCCGG 360
DB |||||
QY 301 CCGGACCGTATTCCTTCGCGCCGCCCTCCCCCGCCGCGCCCGGCCCTCCCTCCCGG 360
DB |||||
QY 361 CAGAGCTCGCCTCCCTCGCCTCAGACTGTTTTGTGTAGCAACGGCAACGGCGCGCGG 420
DB |||||
QY 361 CAGAGCTCGCCTCCCTCGCCTCAGACTGTTTTGTGTAGCAACGGCAACGGCGCGCGG 420
DB |||||
QY 421 TTTTCGGCCCGGCTCCCGCGCGCTCCTTGCTTCGGCGGGCCCTCCCGGCCCTTCGTCGTC 480
DB |||||
QY 421 TTTTCGGCCCGGCTCCCGCGCGCTCCTTGCTTCGGCGGGCCCTCCCGGCCCTTCGTCGTC 480
DB |||||
QY 481 GTCTTCTCCCTTCGCGCAGCCCGGCGCCCTTCGCGCGCGCCCAACCGCGCCCTCCCGG 540
DB |||||
QY 481 GTCTTCTCCCTTCGCGCAGCCCGGCGCCCTTCGCGCGCGCCCAACCGCGCCCTCCCGG 540
DB |||||
QY 541 CTCGGGCGCGGTGCGTCCCGCGCGGTTCGCGGCTCCTTTGGCGCGCCCGCTCCCGG 600
DB |||||
QY 541 CTCGGGCGCGGTGCGTCCCGCGCGGTTCGCGGCTCCTTTGGCGCGCCCGCTCCCGG 600
DB |||||
QY 601 TGTTCGCCCGCGCGCTGCGAGCGGTGTATGGGCCCTCCACCATGTGCTGAAGCCCGAG 660
DB |||||
QY 601 TGTTCGCCCGCGCGCTGCGAGCGGTGTATGGGCCCTCCACCATGTGCTGAAGCCCGAG 660
DB |||||
QY 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
DB |||||
QY 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
DB |||||
QY 721 CAGCGCGCGCGCGCTGCCAATGTCGCGAAGCCCGCGCGCGCGCTTCTAGCGTCG 780
DB |||||
QY 721 CAGCGCGCGCGCGCTGCCAATGTCGCGAAGCCCGCGCGCGCGCTTCTAGCGTCG 780
DB |||||
QY 781 CCGCGCGCGCGCTTCCGCGCTCCTCGCTCCTCGGTCTCCTCGCTCGCGCACGCGCTCCC 840
DB |||||
QY 781 CCGCGCGCGCGCTTCCGCGCTCCTCGCTCCTCGGTCTCCTCGCTCGCGCACGCGCTCCC 840
DB |||||
QY 841 TCCTCGGTGTCGCGCGACCTTCGCGCGCGGAGCGCGCGCTCGGCGAGAGTGCAGAAC 900
DB |||||
QY 841 TCCTCGGTGTCGCGCGACCTTCGCGCGCGGAGCGCGCGCTCGGCGAGAGTGCAGAAC 900
DB |||||
QY 901 AGTACAAAGGACTGCCTCAGCTACGATTTCTTTTGATGGNATCTATGCAATATGAG 960
DB |||||
QY 901 AGTACAAAGGACTGCCTCAGCTACGATTTCTTTTGATGGNATCTATGCAATATGAG 960
DB |||||
QY 961 ATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAATGGA 1020
DB |||||
QY 961 ATGGTTTCATATCTTACATCAGTTGTTGGCTCCAAATGTGAAGTACAAATGGA 1020
DB |||||
QY 1021 GGTATATATGAAGGATTTTAAACTTACAGTCCGAACTGATGTTGGTACTTGATGCC 1080
DB |||||
QY 1021 GGTATATATGAAGGATTTTAAACTTACAGTCCGAACTGATGTTGGTACTTGATGCC 1080
DB |||||
QY 1081 GCACATGAAAGTACAGATCCAGTTCGGGCGGAAACGTGAAGAAATATGAGAGT 1140
DB |||||
QY 1081 GCACATGAAAGTACAGATCCAGTTCGGGCGGAAACGTGAAGAAATATGAGAGT 1140
DB |||||
QY 1141 ATTTTGTTCAAATGTTTCAGACTTTGTTGGTACAGTTTAAAGATATGAGTCCAGTTAT 1200
DB |||||
QY 1141 ATTTTGTTCAAATGTTTCAGACTTTGTTGGTACAGTTTAAAGATATGAGTCCAGTTAT 1200
DB |||||
QY 1201 GCAAAAGAGATGCTTTTATCTGACTCTGCTATCAGTGTAAAGTGAATGGCGAAACAAA 1260
DB |||||
QY 1201 GCAAAAGAGATGCTTTTATCTGACTCTGCTATCAGTGTAAAGTGAATGGCGAAACAAA 1260
DB |||||
QY 1261 GAGAGGACCTGGAGCCTGGGATCGAGTGAACTCAGAGCCCAATGAGAACTTGAGGCT 1320
DB |||||
QY 1261 GAGAGGACCTGGAGCCTGGGATCGAGTGAACTCAGAGCCCAATGAGAACTTGAGGCT 1320
DB |||||
QY 1321 TTGGAAAAATGACGTATCTAAATGGATGGGATCCCAATGATATGTTTCGATATAATGAAGA 1380
DB |||||
QY 1321 TTGGAAAAATGACGTATCTAAATGGATGGGATCCCAATGATATGTTTCGATATAATGAAGA 1380
DB |||||

QY 1381 AATTATGCTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA 1440
DB |||||
QY 1381 AATTATGCTGTAGTGTCTACGTATGATAGCAGTTTATCTTCGTATACAGTCCCTTAGAA 1440
DB |||||
QY 1441 AGAGTAACTCAGAAAGATTTTAAACGGGAAGCAAGGGCAACACAGTTAGCAGAA 1500
DB |||||
QY 1441 AGAGTAACTCAGAAAGATTTTAAACGGGAAGCAAGGGCAACACAGTTAGCAGAA 1500
DB |||||
QY 1501 ATTGAGTCAAGTGCAGTACAAAGCTCAGTGGCCCTGGAAATATGATAGTAGAGT 1560
DB |||||
QY 1501 ATTGAGTCAAGTGCAGTACAAAGCTCAGTGGCCCTGGAAATATGATAGTAGAGT 1560
DB |||||
QY 1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACCTGAGGGGACACAGATAAAC 1620
DB |||||
QY 1561 GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACCTGAGGGGACACAGATAAAC 1620
DB |||||
QY 1621 ACTAGGGAAAAATAAATATATTCCTCTGGACAAAGAAATAGAGAGTCATATCCTGGGA 1680
DB |||||
QY 1621 ACTAGGGAAAAATAAATATATTCCTCTGGACAAAGAAATAGAGAGTCATATCCTGGGA 1680
DB |||||
QY 1681 AGTGGGAGACAGAAATTCACCGCTATGGGCGAGCTGGATCGGCTCCATGCCATCAAGA 1740
DB |||||
QY 1681 AGTGGGAGACAGAAATTCACCGCTATGGGCGAGCTGGATCGGCTCCATGCCATCAAGA 1740
DB |||||
QY 1741 TCCACTTCTCACCTTCAGATTTCAAACCGAAATTCGTTTCAGACCCAAAGAGTAGTTAAT 1800
DB |||||
QY 1741 TCCACTTCTCACCTTCAGATTTCAAACCGAAATTCGTTTCAGACCCAAAGAGTAGTTAAT 1800
DB |||||
QY 1801 GGAGGTGTTCCCTCGCCCATCGCTTGCCCATCTCTCTCTCGCCCATCTCTCGCTAC 1860
DB |||||
QY 1801 GGAGGTGTTCCCTCGCCCATCGCTTGCCCATCTCTCTCTCGCCCATCTCTCGCTAC 1860
DB |||||
QY 1861 CAGTCAGTCCCAACTCTCTTCCACCTCGGCGAGCACCCCTACACGGCGCCCTCCAGG 1920
DB |||||
QY 1861 CAGTCAGTCCCAACTCTCTTCCACCTCGGCGAGCACCCCTACACGGCGCCCTCCAGG 1920
DB |||||
QY 1921 CCCCCTCGCGGCGCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTCTCCAGTCTCT 1980
DB |||||
QY 1921 CCCCCTCGCGGCGCATCCAGACCCCGTCTCACCCCTCTGCTCATGTTCTCCAGTCTCT 1980
DB |||||
QY 1981 GTCTCTACTATGCTTAAACGCATGCTTTCAGAAAGGCTCCAAAGATGTCCTCCAAAGGCC 2040
DB |||||
QY 1981 GTCTCTACTATGCTTAAACGCATGCTTTCAGAAAGGCTCCAAAGATGTCCTCCAAAGGCC 2040
DB |||||
QY 2041 CAGCGACATCTCGAAATTCACAGAGTTTCTGCTGGAGGGGTTCATATCCAGTGGCCTA 2100
DB |||||
QY 2041 CAGCGACATCTCGAAATTCACAGAGTTTCTGCTGGAGGGGTTCATATCCAGTGGCCTA 2100
DB |||||
QY 2101 GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGACGAGT 2160
DB |||||
QY 2101 GAATTTGTATCCCAACCCACCCAGTGAAGCAGTACTCTCCAGTAGCAAGGACGAGT 2160
DB |||||
QY 2161 CCCCCTGGGGGAAACGTGCTCATCAGTGGTTCAGAGGGGTTCAGAGTTATCCCTTAAACT 2220
DB |||||
QY 2161 CCCCCTGGGGGAAACGTGCTCATCAGTGGTTCAGAGGGGTTCAGAGTTATCCCTTAAACT 2220
DB |||||
QY 2221 CATAGACCCAGGTCTCCAGACAGAAACAGATTTGGAAATACCCCAGTGGGCGAGTCTT 2280
DB |||||
QY 2221 CATAGACCCAGGTCTCCAGACAGAAACAGATTTGGAAATACCCCAGTGGGCGAGTCTT 2280
DB |||||
QY 2281 GCTTCTCCCCAAGCTGTTATTTCCAACTGAAGCTGTTGCGCATGCTATTTCCAGTCTGA 2340
DB |||||
QY 2341 TCTCCTAGCCTGCTAGTCTCGCATCGAAACAGAGCTGTTTACCCCTTCTAGTGAGGCTAAA 2400
DB |||||
QY 2341 TCTCCTAGCCTGCTAGTCTCGCATCGAAACAGAGCTGTTTACCCCTTCTAGTGAGGCTAAA 2400
DB |||||
QY 2401 GATTCCAGGCTTCAAGATCAGAGGACAGAACTCTCTCGCAGGGAATAAGAAAAATTTAAA 2460
DB |||||
QY 2401 GATTCCAGGCTTCAAGATCAGAGGACAGAACTCTCTCGCAGGGAATAAGAAAAATTTAAA 2460
DB |||||
QY 2461 CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAAAACAAAGGTATATATCACAGTGTGT 2520
DB |||||

2461 DB |||||CCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAACCAAAAGGTATATCAACAGTGT 2520
2521 QY TCTGAAATAGAGAAAACAGATTGATGATTTAAAGAAATTTAAGATGATTTTAGTTACAG 2580
2521 DB TCTGAAATAGAGAAAACAGATTGATGATTTAAAGAAATTTAAGATGATTTTAGTTACAG 2580
2581 QY CCAAGTTCTACTTCTGAAATCTATGATCAACTACTAAACCAAAATAGAGAGGAGAAA 2640
2581 DB CCAAGTTCTACTTCTGAAATCTATGATCAACTACTAAACCAAAATAGAGAGGAGAAA 2640
2641 QY TCAAGAGATTTGATCAAGACAAAATTTGAACCAAGTCTTAAGGATTTCTTTCAATGAAAAT 2700
2641 DB TCAAGAGATTTGATCAAGACAAAATTTGAACCAAGTCTTAAGGATTTCTTTCAATGAAAAT 2700
2701 QY AGCAGCAGCAACTGTACCAAGTGGCAGCAGCAAGCCGAATAGCCCAAGCAATTTCCCTTCA 2760
2701 DB AGCAGCAGCAACTGTACCAAGTGGCAGCAGCAAGCCGAATAGCCCAAGCAATTTCCCTTCA 2760
2761 QY ATACTTTAGTAAACACGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCAAGCT 2820
2761 DB ATACTTTAGTAAACACGAGCACAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCAAGCT 2820
2821 QY TCCAGCCAGCATGTAAAACAGAGAAAGACGATTAAGGAAAGAGAAAGACGACGCTGAG 2880
2821 DB TCCAGCCAGCATGTAAAACAGAGAAAGACGATTAAGGAAAGAGAAAGACGACGCTGAG 2880
2881 QY CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCAAGTTCTTCTCT 2940
2881 DB CAAAGTTAGGAAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCAAGTTCTTCTCT 2940
2941 QY CAGCCAAAGCTTCTACTACCCCAACTTCACTCGGCCTCAAGCAACAACCTAGCCCATCT 3000
2941 DB CAGCCAAAGCTTCTACTACCCCAACTTCACTCGGCCTCAAGCAACAACCTAGCCCATCT 3000
3001 QY ATGGTGGGTCAACAGCCAACTCAGTTTATCTCAGGCTGTTTGTGTCACCAAT 3060
3001 DB ATGGTGGGTCAACAGCCAACTCAGTTTATCTCAGGCTGTTTGTGTCACCAAT 3060
3061 QY ATGATGTATCCAGTCCAGTGAGCCAGGCTGCAACCTTTATACCAATACCTATGACG 3120
3061 DB ATGATGTATCCAGTCCAGTGAGCCAGGCTGCAACCTTTATACCAATACCTATGACG 3120
3121 QY CCAATGCCAGTGAATCAAGCAAGACATATAGAGCAGTACCAATATGCCCAACAGCGG 3180
3121 DB CCAATGCCAGTGAATCAAGCAAGACATATAGAGCAGTACCAATATGCCCAACAGCGG 3180
3181 QY CAAGACAGCATCATCAGAGTGCATGATGCAACCCAGGTCAGAGGGGCCCAACCGATT 3240
3181 DB CAAGACAGCATCATCAGAGTGCATGATGCAACCCAGGTCAGAGGGGCCCAACCGATT 3240
3241 QY GCAGCCACCCAGCTTACTCCAGCAATATGTCCTACAGTCTCCTCAGAGGTTCCCA 3300
3241 DB GCAGCCACCCAGCTTACTCCAGCAATATGTCCTACAGTCTCCTCAGAGGTTCCCA 3300
3301 QY AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCTATGTCTATGT 3360
3301 DB AATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCTATGTCTATGT 3360
3361 QY CCTGTAATACAGGGTAATGTAGAAATGATGGCAACCAACACACGCGCCAGCTGGTTTA 3420
3361 DB CCTGTAATACAGGGTAATGTAGAAATGATGGCAACCAACACACGCGCCAGCTGGTTTA 3420
3421 QY GTATCTTCTTTCAGCACTCAGTACGGGCTCATGAGCAGACGATGCGATGATGATGT 3480
3421 DB GTATCTTCTTTCAGCACTCAGTACGGGCTCATGAGCAGACGATGCGATGATGATGT 3480
3481 QY CCAAAATACCAACAAAGGAGACAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3540
3481 DB CCAAAATACCAACAAAGGAGACAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3540
3541 QY TCCCTTGCTCAGAGTATGCGACCCCTTAACGCTACCTGCAACCCACATCTCCACACCT 3600

3541 DB TCCCTTGCTCAGCAGTATGCGCACCCCTAACGCTACCTTGCACCCACATACTCCACACCT 3600
3601 QY CAGCCTTCAGCTTACCCCACTGGACAGAGCAAAACCAACATGTTGGAGTCTATCTTGA 3660
3601 DB CAGCCTTCAGCTTACCCCACTGGACAGAGCAAAACCAACATGTTGGAGTCTATCTTGA 3660
3661 QY CCAAGTCTCTGTTTCAGCATCAGCACGAGCGGCCAGGCTCTCCATCTGCGCAGTCCA 3720
3661 DB CCAAGTCTCTGTTTCAGCATCAGCACGAGCGGCCAGGCTCTCCATCTGCGCAGTCCA 3720
3721 QY CAGCAGCAGTCTCAGCATTTTACCAACGCGGGCTTGGCCCAACTCCACCTTCCATGACACCT 3780
3721 DB CAGCAGCAGTCTCAGCATTTTACCAACGCGGGCTTGGCCCAACTCCACCTTCCATGACACCT 3780
3781 QY GCCTCCAAACAGCAGTCTGCGCACAGAAATGTTTCCAGCAGCAACAACAGACTGTCTTTAG 3840
3781 DB GCCTCCAAACAGCAGTCTGCGCACAGAAATGTTTCCAGCAGCAACAACAGACTGTCTTTAG 3840
3841 QY ATCCATCTCTCTCAGCTTTCAGCGCGGTATACCAACCCACCCACATGCGCCACGTTACCT 3900
3841 DB ATCCATCTCTCTCAGCTTTCAGCGCGGTATACCAACCCACCCACATGCGCCACGTTACCT 3900
3901 QY CAGGCTCATGTACAGTCTCAGGAATGTTCTTCTCATCCAACTGCCCATCGCCCAATGATG 3960
3901 DB CAGGCTCATGTACAGTCTCAGGAATGTTCTTCTCATCCAACTGCCCATCGCCCAATGATG 3960
3961 QY CTAATGAACACACAGCCAGCCCGGCTCCAGGCGCCCTCGCTCAAAAGTGCACCTACAG 4020
3961 DB CTAATGAACACACAGCCAGCCCGGCTCCAGGCGCCCTCGCTCAAAAGTGCACCTACAG 4020
4021 QY CCAATTCAGTCTCAGCAACAGCGCATTTTCCCTTATATGACGACCCCTTTCAGTACAAGCC 4080
4021 DB CCAATTCAGTCTCAGCAACAGCGCATTTTCCCTTATATGACGACCCCTTTCAGTACAAGCC 4080
4081 QY CACCAACACAGCAGTTGTAAGCTGCGCTGGAGAAACGAAAGGCCAAATTCCTTCTCTC 4140
4081 DB CACCAACACAGCAGTTGTAAGCTGCGCTGGAGAAACGAAAGGCCAAATTCCTTCTCTC 4140
4141 QY CTTCTTACTGCTTCTACCAACTGGAAGCAGACAGAAAATAGAAATTTTATTTTGT 4200
4141 DB CTTCTTACTGCTTCTACCAACTGGAAGCAGACAGAAAATAGAAATTTTATTTTGT 4200
4201 QY TAAATATATATGTTGATTTCTTGTAACTCCAATAGGAATGCTAAGTCTTCACTTGCAG 4260
4201 DB TAAATATATATGTTGATTTCTTGTAACTCCAATAGGAATGCTAAGTCTTCACTTGCAG 4260
4261 QY TGAAGATATCTTGGACCGAGTAGAGCATTTAGAACTTTGGGGCTTATTCATAATTTCCA 4320
4261 DB TGAAGATATCTTGGACCGAGTAGAGCATTTAGAACTTTGGGGCTTATTCATAATTTCCA 4320
4321 QY TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGTTGCCGAAACTGGAAATTTATTT 4380
4321 DB TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGTTGCCGAAACTGGAAATTTATTT 4380
4381 QY ATTTTAAATAACCTTGAAGTCATGAACATCAGTCTAGCAAAAGAAAGTACAAGAT 4440
4381 DB ATTTTAAATAACCTTGAAGTCATGAACATCAGTCTAGCAAAAGAAAGTACAAGAT 4440
4441 QY GATTCTTGCTGCTTACTCTTAAATAAAAAAAAAAAAAAAAAAAAA 4481
4441 DB GATTCTTGCTGCTTACTGTAAAAAAAAAAAAAAAAAAAAA 4481

RESULT 2

US-10-750-323-2

; Sequence 2, Application US/10750323

; Publication No. US20050032083A1

; GENERAL INFORMATION:

; APPLICANT: Pulst, Stefan M

; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

ADDRESSEE: Muetting, Raasch & Gebhardt, P.A.

STREET: 119 North Fourth Street

CITY: Minneapolis

STATE: Minnesota

COUNTRY: USA

ZIP: 55401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/750,323

FILING DATE: 30-Dec-2003

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/727,084

FILING DATE: 08-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Muetting, Ann M.

REGISTRATION NUMBER: 33,977

REFERENCE/DOCKET NUMBER: 232.00010101

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612/305-1220

TELEFAX: 612/305-1228

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 4481 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 163..4101

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-750-323-2

Query Match 100.0%; Score 4481; DB 8; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCCTCAGTGTCTCTCCCGCGCGCGCGG 60

1 ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCCTCAGTGTCTCTCCCGCGCGCGG 60

61 GCACCTCAGCTTCGCTTCGCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120

61 GCACCTCAGCTTCGCTTCGCTGACCCCTCCGACTTCGCGTAAAGAGTCCCTATCCG 120

121 CACCTCCGCTCCACCCGCGCTCGCGCGCGCGCGCTCCGATCGCTCAGCGCGCGCA 180

121 CACCTCCGCTCCACCCGCGCTCGCGCGCGCGCGCTCCGATCGCTCAGCGCGCGCA 180

181 GCTCTCGAGTCCCGGCTGCGCAACCGAGTTCGCGTTCGCGTAAAGAGTCCCTATCCG 240

181 GCTCTCGAGTCCCGGCTGCGCAACCGAGTTCGCGTTCGCGTAAAGAGTCCCTATCCG 240

241 GGGTGGCGCTCGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300

241 GGGTGGCGCTCGCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300

301 CCGGACCGTATCCCTCCGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360

301 CCGGACCGTATCCCTCCGCGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 360

361 CAGACTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACCGGCAACCGGCGCGCGCG 420

361 CAGACTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAACCGGCAACCGGCGCGCGCG 420

421 TTTCCGCGCGCTCCCGCGCGCTTCGCTCGCGCGCGCTCCCGCGCGCGCTTCGCTCG 480

421 TTTCCGCGCGCTCCCGCGCGCTTCGCTCGCGCGCGCTTCGCTCGCGCGCGCTTCGCTCG 480

481 GTCCTTCTCCCTCGCCAGCCCGGGCGCCCTCCTCGGCGCGCGCAACCCGCGCTCCCGC 540

481 GTCCTTCTCCCTCGCCAGCCCGGGCGCCCTCCTCGGCGCGCGCAACCCGCGCTCCCGC 540

541 CTCGGCGCGTGGTCCCGCGCGTTCGGGCTCTCTTGGCGCGCGCGCTCCCGC 600

541 CTCGGCGCGTGGTCCCGCGCGTTCGGGCTCTCTTGGCGCGCGCGCTCCCGC 600

601 TGTCTCCCGCGCGTGCAGCCGCTGTATGGGCGCCCTCACCATGTCTCGCTCAAGCCCCAG 660

601 TGTCTCCCGCGCGTGCAGCCGCTGTATGGGCGCCCTCACCATGTCTCGCTCAAGCCCCAG 660

661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 720

661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAG 720

721 CAGCGCGCGCGCGCTGCCAATGTCCGCAAGCCCGCGCGCGCGCTTCTAGCGTCG 780

721 CAGCGCGCGCGCGCTGCCAATGTCCGCAAGCCCGCGCGCGCGCTTCTAGCGTCG 780

781 CCGCGCGCGCGCTTCCCGCTCCTCGTCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCT 840

781 CCGCGCGCGCGCTTCCCGCTCCTCGTCTCTCGTCTCTCGTCTCTCGTCTCTCGTCTCT 840

841 TCCTCGTGTGGTGGCGGACCTCCGCGCGCGGAGCGCGCGCTGGCAGAGTTCGAAAC 900

841 TCCTCGTGTGGTGGCGGACCTCCGCGCGCGGAGCGCGCGCTGGCAGAGTTCGAAAC 900

901 AGTAACAAAGACTGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAAAATATGAGG 960

901 AGTAACAAAGACTGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAAAATATGAGG 960

961 ATGGTTCAATATACATCAGTTGTGGCTCCAAATGTGAAGTACAAATGGAATGGA 1020

961 ATGGTTCAATATACATCAGTTGTGGCTCCAAATGTGAAGTACAAATGGAATGGA 1020

1021 GGTATATGAAGAGTGTAAACTTACAGTCCGAAAGTGTGATTTGGTATCTTGTATGCC 1080

1021 GGTATATGAAGAGTGTAAACTTACAGTCCGAAAGTGTGATTTGGTATCTTGTATGCC 1080

1081 GCATGAGAAAGTACAGAAATCCAGTTCGGGGCGGAAACGTAAGAAATATGGAGAGT 1140

1081 GCATGAGAAAGTACAGAAATCCAGTTCGGGGCGGAAACGTAAGAAATATGGAGAGT 1140

1141 ATTTGTTCAATGTTCAGACTTTTGTGTGTACAGTTTAAAGATATCGACTCCAGTTAT 1200

1141 ATTTGTTCAATGTTCAGACTTTTGTGTGTGTACAGTTTAAAGATATCGACTCCAGTTAT 1200

1201 GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCAAGTGAATGGCGAACACAAA 1260

1201 GCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCAAGTGAATGGCGAACACAAA 1260

1261 GAGAAGGACCTGGAGCCCTGGGATGCGAGTGAATCAGCAGCAATGAGGAACTTGGAGCT 1320

1261 GAGAAGGACCTGGAGCCCTGGGATGCGAGTGAATCAGCAGCAATGAGGAACTTGGAGCT 1320

1321 TTGAAATATGAGTATCTAAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA 1380

1321 TTGAAATATGAGTATCTAAATGGATGGATCCCAATGATATGTTTCGATATATGAAGAA 1380

1381 AATTATGTTAGTGTCTACGTATGATAGAGTTTATCTTCGTATACAGTCCCTTTAGAA 1440

1381 AATTATGTTAGTGTCTACGTATGATAGAGTTTATCTTCGTATACAGTCCCTTTAGAA 1440

1441 AGAGATAACTCAGAAAGATTTTAAACGGGAAGCAAGGGCAAAACCAAGTTAGCAGAAAG 1500

1441 AGAGATAACTCAGAAAGATTTTAAACGGGAAGCAAGGGCAAAACCAAGTTAGCAGAAAG 1500

1501 ATTGAGTCAAGTGGCGCGTCAAGAGCTCGAGTGGCGCGCGCTTGGGAAATGATAGGAGT 1560

1501 ATTGAGTCAAGTGGCGCGTCAAGAGCTCGAGTGGCGCGCGCTTGGGAAATGATAGGAGT 1560

QY	1561	GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACTGTCAGGGGACAGCATAAAC	1620
DB	1561	GAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACTGTCAGGGGACAGCATAAAC	1620
QY	1621	ACTAGGGAAAAATAATATATTCCTCTGGACAAAGAAATAGAGAAGTCATATCTCTGGGA	1680
DB	1621	ACTAGGGAAAAATAATATATTCCTCTGGACAAAGAAATAGAGAAGTCATATCTCTGGGA	1680
QY	1681	AGTGGGAGACAGAAATTCACCGCTGATGGCCAGCGCTGGATCGGGCTCCATGCCATCAAGA	1740
DB	1681	AGTGGGAGACAGAAATTCACCGCTGATGGCCAGCGCTGGATCGGGCTCCATGCCATCAAGA	1740
QY	1741	TCCACTTCTCACATTCAGATTTCCAAACCGAAATTCGGTTTCAGACCAAGAGTAGTTAAT	1800
DB	1741	TCCACTTCTCACATTCAGATTTTCGTTCTCTGGACAAAGAAATAGAGAAGTCATATCTCTGGGA	1800
QY	1801	GGAGGTGTTCCTGGCCATCGCTTGCCCATCTCTCTCTGCTCGCCACCTCTCTCGCTAC	1860
DB	1801	GGAGGTGTTCCTGGCCATCGCTTGCCCATCTCTCTCTCTGCTCGCCACCTCTCTCGCTAC	1860
QY	1861	CAGTCAGGTCCCAACTCTCTCTCCACTCGGGCAGCCACCCCTCACAGGCCCGCCCTCCAGG	1920
DB	1861	CAGTCAGGTCCCAACTCTCTCTCCACTCGGGCAGCCACCCCTCACAGGCCCGCCCTCCAGG	1920
QY	1921	CCCCCTCGCGCCATTCAGACACCCCGTCTCACCCCTCTGCTCANGTTCCTCAGTCTCT	1980
DB	1921	CCCCCTCGCGCCATTCAGACACCCCGTCTCACCCCTCTGCTCANGTTCCTCAGTCTCT	1980
QY	1981	GTCTCTACTATGCTTAAACGATGTCCTTCAGAAAGGCTCCAGGATGTCCTCCAAAGGCC	2040
DB	1981	GTCTCTACTATGCTTAAACGATGTCCTTCAGAAAGGCTCCAGGATGTCCTCCAAAGGCC	2040
QY	2041	CAGCGACATCTCTCGAAATTCACAGAGTTTCTGCTGGAGGGGTTCATATCCAGTGGCCTA	2100
DB	2041	CAGCGACATCTCTCGAAATTCACAGAGTTTCTGCTGGAGGGGTTCATATCCAGTGGCCTA	2100
QY	2101	GAATTGTATCCCAACACCCAGTCAGTGAAGAGTACTCTCTCAGTAGCAAGGACCACT	2160
DB	2101	GAATTGTATCCCAACACCCAGTCAGTGAAGAGTACTCTCTCAGTAGCAAGGACCACT	2160
QY	2161	CCCTCGGGGGAACTGTGTTCATCAGTGTTCAGTGGGTTCAGAGTTATCCCTTAAACT	2220
DB	2161	CCCTCGGGGGAACTGTGTTCATCAGTGTTCAGTGGGTTCAGAGTTATCCCTTAAACT	2220
QY	2221	CATAGACCCAGTCTCCAGACAGAACTAGTTGGAAATACCCAGTGGGCGAGTCTCT	2280
DB	2221	CATAGACCCAGTCTCCAGACAGAACTAGTTGGAAATACCCAGTGGGCGAGTCTCT	2280
QY	2281	GCTTCTCCCAAGCTGGTATTAATTCACACTGAAGCTGTGGCATGCTATTCAGCTGCA	2340
DB	2281	GCTTCTCCCAAGCTGGTATTAATTCACACTGAAGCTGTGGCATGCTATTCAGCTGCA	2340
QY	2341	TCTCTACGCTGTAGTCTCGATTCGAACAGAGCTGTACCCCTTCTAGTGGGCTAAA	2400
DB	2341	TCTCTACGCTGTAGTCTCGATTCGAACAGAGCTGTACCCCTTCTAGTGGGCTAAA	2400
QY	2401	GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAGAAATATTAAA	2460
DB	2401	GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAGAAATATTAAA	2460
QY	2461	CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAAAAACAAGGTATATCACCAAGTTGTT	2520
DB	2461	CCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAAAAACAAGGTATATCACCAAGTTGTT	2520
QY	2521	TCTGAACATAGAAAAACAGATTGATTTAAAGAAATTTAAGATGATTTTAGGTTACAG	2580
DB	2521	TCTGAACATAGAAAAACAGATTGATTTAAAGAAATTTAAGATGATTTTAGGTTACAG	2580
QY	2581	CCAAGTTCTACTTCTCAATCTATGATCAACTACTTAAACAAAAATAGAGGGGAGAAAAA	2640
DB	2581	CCAAGTTCTACTTCTCAATCTATGATCAACTACTTAAACAAAAATAGAGGGGAGAAAAA	2640
QY	2641	TCAGAGATTTGATTCAAAGACAAAAATGAAACCAAGTGTGAAGATTTCTTTTCAATGAAAT	2700

Db	2641	TCAAGAGATTGTGATCAAAAGACAAAATTTGAACCAAGTGTAAAGATTCTTTTCATTGAAAT2700
Qy	2701	AGCAGCAGCAACTGTGTACCAAGTGGCAGCAGCAAGCCGAATAGCCCCAGCAATTTCCCTTCA2760
Db	2701	AGCAGCAGCAACTGTGTACCAAGTGGCAGCAGCAAGCCGAATAGCCCCAGCAATTTCCCTTCA2760
Qy	2761	ATACTTAGTAAACAGCAGGAGCAAAAGAGGGACCTGAGGTCACATTCCTCCAAGGGTTACAGCT2820
Db	2761	ATACTTAGTAAACAGCAGGAGCAAAAGAGGGACCTGAGGTCACATTCCTCCAAGGGTTACAGCT2820
Qy	2821	TCCAGCCCCAGCATGTATAAACAAGAGAAGAAGCATTAAGCAAGAGAAGAAGACGCGAGCTGAG2880
Db	2821	TCCAGCCCCAGCATGTATAAACAAGAGAAGAAGCATTAAGCAAGAGAAGAAGACGCGAGCTGAG2880
Qy	2881	CAAGTTAGGAAATCAACATTGTAAATCCCAATGCAAAAGGAGTTCAACCCAGCTTCTTCTCT2940
Db	2881	CAAGTTAGGAAATCAACATTGTAAATCCCAATGCAAAAGGAGTTCAACCCAGCTTCTTCTCT2940
Qy	2941	CAGCAAGCCTTCTACTAGCCCAATCTTACCTCGGCCCTCAAGCACAACCTAGACCCATCT3000
Db	2941	CAGCAAGCCTTCTACTAGCCCAATCTTACCTCGGCCCTCAAGCACAACCTAGACCCATCT3000
Qy	3001	ATGGTGGGTATCAACAGCACAATCCAGTTTATCTACGCTGTTTGTGTTGACCAAAAT3060
Db	3001	ATGGTGGGTATCAACAGCACAATCCAGTTTATCTACGCTGTTTGTGTTGACCAAAAT3060
Qy	3061	ATGATGTATCCAGTCCAGTCGAGCCAGGCGTGCAACCTTTATACCCAATACCTATGACG3120
Db	3061	ATGATGTATCCAGTCCAGTCGAGCCAGGCGTGCAACCTTTATACCCAATACCTATGACG3120
Qy	3121	CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAATATGCCCCAACAGCGG3180
Db	3121	CCCATGCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCAATATGCCCCAACAGCGG3180
Qy	3181	CAGACCAGCATCATCAGAGTGCCATGATGSCACCCAGCGTCAGCAGGGGCCACCGATT3240
Db	3181	CAGACCAGCATCATCAGAGTGCCATGATGSCACCCAGCGTCAGCAGGGGCCACCGATT3240
Qy	3241	GCAGCCACCCACAGCTTACTCCAGCCAATATGTTGCTTACAGTCCTCAGCAGTTTCCCA3300
Db	3241	GCAGCCACCCACAGCTTACTCCAGCCAATATGTTGCTTACAGTCCTCAGCAGTTTCCCA3300
Qy	3301	AATCAGCCCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGT3360
Db	3301	AATCAGCCCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGT3360
Qy	3361	CCTGTAAATACAGGTAAATGCTAGAAATGATGGCACCCACCAACACGCCCGAGCTGGTTTA3420
Db	3361	CCTGTAAATACAGGTAAATGCTAGAAATGATGGCACCCACCAACACGCCCGAGCTGGTTTA3420
Qy	3421	GTATCTTTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCGATGGATGTATGCAATG3480
Db	3421	GTATCTTTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCGATGGATGTATGCAATG3480
Qy	3481	CCCAAAATTAACATCAAACAAGGAGACAAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT3540
Db	3481	CCCAAAATTAACATCAAACAAGGAGACAAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT3540
Qy	3541	TCCCTTGTTCAGCATATGCGCACCCTTAAACGCTACCTCTGCACCCACATCTCCACACCT3600
Db	3541	TCCCTTGTTCAGCATATGCGCACCCTTAAACGCTACCTCTGCACCCACATCTCCACACCT3600
Qy	3601	CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCCAAATGTTGGAAGTCACTCTGCA3660
Db	3601	CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCCAAATGTTGGAAGTCACTCTGCA3660
Qy	3661	CCCAGTCTCTTTCAGCACCATCAGCAGCAGGCGGCCAGGCTCTCCATCTGGCGAGTCCA3720
Db	3661	CCCAGTCTCTTTCAGCACCATCAGCAGCAGGCGGCCAGGCTCTCCATCTGGCGAGTCCA3720
Qy	3721	CAGCAGCATCAGCCATTTACCGCGGGGCTTGGCGCAACTCCACCTCCATGACACCT3780

Db 3721 CAGCAGCAGTACGCCATTTACACGGGGGCTTGGCCCACTCCACCTCCATGACACT 3780
Qy 3781 GCCTCCAAACACGAGTCGCGACAGAAATAGTTTCCAGCAGCACACAGACTGCTTTTACG 3840
Db 3781 GCCTCCAAACACGAGTCGCGACAGAAATAGTTTCCAGCAGCACACAGACTGCTTTTACG 3840
Qy 3841 ATCCATCTCTTCACAGTTTCAGCGGGGTATACCAACCCACATGATGCCCACTACCT 3900
Db 3841 ATCCATCTCTCTTCAGGTTTCAGCGGGGTATACCAACCCACATGATGCCCACTACCT 3900
Qy 3901 CAGGCTCATGTACAGTCAGGAATGTTCTCTCATCCAACTGCCCATGCGGCAATGATG 3960
Db 3901 CAGGCTCATGTACAGTCAGGAATGTTCTCTCATCCAACTGCCCATGCGGCAATGATG 3960
Qy 3961 CTAATGAGGACACAGCCACCGGGGCTGCCAGGCGGCCCTCGCTCAAAGTGCACTACAG 4020
Db 3961 CTAATGAGGACACAGCCACCGGGGCTGCCAGGCGGCCCTCGCTCAAAGTGCACTACAG 4020
Qy 4021 CCATTTCCAGTCTCGACAAACAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAGCC 4080
Db 4021 CCATTTCCAGTCTCGACAAACAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAGCC 4080
Qy 4081 CACCACCAACAGCAGTTGTAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCTCTC 4140
Db 4081 CACCACCAACAGCAGTTGTAGGCTGCCCTGGAGGAACCGAAAGGCCAAATTCCTCTCTC 4140
Qy 4141 CTTCTACTGCTTTACCAACTGGAAGCACAGAAACTAGAAATTTCAATTTTGTGTTTT 4200
Db 4141 CTTCTACTGCTTTACCAACTGGAAGCACAGAAACTAGAAATTTCAATTTTGTGTTTT 4200
Qy 4201 TAAATATATATGTTGATTTCTTTGAATCAATCAATAGGAATGCTAACAGTTCACTTGCAG 4260
Db 4201 TAAATATATATGTTGATTTCTTTGAATCAATCAATAGGAATGCTAACAGTTCACTTGCAG 4260
Qy 4261 TGAAGATACCTTGGACCGAGTAGGCGCATTTAGGAATTTGGGGGCTATTCCATAATTTCCA 4320
Db 4261 TGAAGATACCTTGGACCGAGTAGGCGCATTTAGGAATTTGGGGGCTATTCCATAATTTCCA 4320
Qy 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGGCCGAACTGGAAGTTATTT 4380
Db 4321 TATGCTGTTTTCAGAGTCCCGCAGGTACCCAGCTCTGCTTGGCCGAACTGGAAGTTATTT 4380
Qy 4381 ATTTTAAATACCTTTGAAAGTCATGAAACATCATAGCTAGCAAAAGAAAGTAAACAGAGT 4440
Db 4381 ATTTTAAATACCTTTGAAAGTCATGAAACATCATAGCTAGCAAAAGAAAGTAAACAGAGT 4440
Qy 4441 GATTCCTGCTGCTATTACTGCTTAAAAAATAAAAAA 4481
Db 4441 GATTCCTGCTGCTATTACTGCTTAAAAAATAAAAAA 4481

RESULT 3
US-10-887-553A-493
; Sequence 493, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
; APPLICANT: Garza, Dan
; APPLICANT: Li, Hao
; TITLE OF INVENTION: Method to treat conditions associated
; with insulin signalling dysregulation
; FILE OF INVENTION: 4-33262
; CURRENT APPLICATION NUMBER: US/10/887,553A
; CURRENT FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/485,883
; PRIOR FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 1208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: human
US-10-887-553A-493

Query Match 100.0%; Score 4481; DB 9; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTGTCCCTCCGGCCCCCGG 60
Db 1 ACCCCGAGAAAGCAACCCAGCGCGCGCGCTCTCACGTGTCCCTCCGGCCCCCGG 60
Qy 61 GCCACCTCAGCTTCTGCTTCCGCTCTGACCCCTCCAGACTTCCGCTAAAGAGTCCCTATCCG 120
Db 61 GCCACCTCAGCTTCTGCTTCCGCTCTGACCCCTCCAGACTTCCGCTAAAGAGTCCCTATCCG 120
Qy 121 CACCTCCGCTCCCAACCCGCGGCTCTCGGCGCGCCCTCCGATGCGCTCAGCGCGCGCA 180
Db 121 CACCTCCGCTCCCAACCCGCGGCTCTCGGCGCGCCCTCCGATGCGCTCAGCGCGCGCA 180
Qy 181 GCTCCTCGGAGTCCCGGCTGCGCACCGAGTCTCGCGCTTCGCGCGAGCGAGTGGGCC 240
Db 181 GCTCCTCGGAGTCCCGGCTGCGCACCGAGTCTCGCGCTTCGCGCGAGCGAGTGGGCC 240
Qy 241 GGGTGGGCTCGCTCCAGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 241 GGGTGGGCTCGCTCCAGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Qy 301 CCGGAGCCGTATCCCTCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 CCGGAGCCGTATCCCTCGCGCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy 361 CAGAGCTCGCTCCCTCGCGCTCAGACTGTTTGGTAGCAACGAGCGCGCGCGCGG 420
Db 361 CAGAGCTCGCTCCCTCGCGCTCAGACTGTTTGGTAGCAACGAGCGCGCGCGCGG 420
Qy 421 TTTCCGCGCGCTCCCGCGCTCTTGGTCTCGCGCGCGCTCCCGCGCGCTTCGTCTGTC 480
Db 421 TTTCCGCGCGCTCCCGCGCTCTTGGTCTCGCGCGCGCTCCCGCGCGCTTCGTCTGTC 480
Qy 481 GTCTTTCTCCCTCGCGCGCTTCCCGCGCGCTTCCCGCGCGCTTCCCGCGCTTCCCG 540
Db 481 GTCTTTCTCCCTCGCGCGCTTCCCGCGCGCTTCCCGCGCGCTTCCCGCGCTTCCCG 540
Qy 541 CTCGCGCGCGTGGTCTCCCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCCCG 600
Db 541 CTCGCGCGCGTGGTCTCCCGCGCTTCCCGCGCTTCCCGCGCTTCCCGCGCTTCCCG 600
Qy 601 TGTCTCCCGCGCGCTCGAGCTGTATGGGCGCTTCCAGCTGCTTCCAGTCCAGCG 660
Db 601 TGTCTCCCGCGCGCTCGAGCTGTATGGGCGCTTCCAGCTGCTTCCAGTCCAGCG 660
Qy 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Db 661 CAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720
Qy 721 CAGCGCGCGCGCTTCCAGTCTGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 721 CAGCGCGCGCGCTTCCAGTCTGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Qy 781 CCGCGCGCGCGCTTCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
Db 781 CCGCGCGCGCGCTTCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 840
Qy 841 TCCTCGGTGTCGCGGACCTCCCGCGCGGAGCGCGCGCTTGGGCGAGGTCGAAAC 900
Db 841 TCCTCGGTGTCGCGGACCTCCCGCGCGGAGCGCGCGCTTGGGCGAGGTCGAAAC 900
Qy 901 AGTAACAAAGAGCTCGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAATATGAGG 960
Db 901 AGTAACAAAGAGCTCGCTCAGTCTACGATTTCTTTGATGGAATCTATGCAATATGAGG 960
Qy 961 ATGTTTCATATATACATCAGTTGTTGGCTCCAAATGTGAAGTACAAATGGA 1020
Db 961 ATGTTTCATATATACATCAGTTGTTGGCTCCAAATGTGAAGTACAAATGGA 1020
Qy 1021 GGTATATAGGAGGAGTGTATGTAAGTCCGAGTGTGATTTGGTACTTGTATGCC 1080

|||||
1021 GGTATATATGAAGGAGTTTTTAAATCTTACAGTCGGAAGTGATTTGGTACTTGATGCC 1080
QY
1081 GCACATGAGAAAGTACAGAACTCCAGTTCCGGGGCGAAACGTTGAAGAAATATGAGAGT 1140
Db
1081 GCACATGAGAAAGTACAGAACTCCAGTTCCGGGGCGAAACGTTGAAGAAATATGAGAGT 1140
QY
1141 ATTTTGTTCAAATGTTTCAGACTTTTGTGTGGTACAGTTTAAAGATATGGAATCCAGTTAT 1200
Db
1141 ATTTTGTTCAAATGTTTCAGACTTTTGTGTGGTACAGTTTAAAGATATGGAATCCAGTTAT 1200
QY
1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGCTTAAAGTGAATGGCCGAAACACAAA 1260
Db
1201 GCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTGCTTAAAGTGAATGGCCGAAACACAAA 1260
QY
1261 GAGAAGGACCTGGAGCCCTGGGATGAGGTGAATCTCAGCCCAATGAGGAACTTGAGGCT 1320
Db
1261 GAGAAGGACCTGGAGCCCTGGGATGAGGTGAATCTCAGCCCAATGAGGAACTTGAGGCT 1320
QY
1321 TTGGAAAAATGACGTATCTAATGGAATGGGATCCCAATGATGTTTCGATATATGAGAA 1380
Db
1321 TTGGAAAAATGACGTATCTAATGGAATGGGATCCCAATGATGTTTCGATATATGAGAA 1380
QY
1381 AATTATGTTGATGCTTACGTATGATAGCAGTTTATCTTCTGATACAGTCCCTTAGAA 1440
Db
1381 AATTATGTTGATGCTTACGTATGATAGCAGTTTATCTTCTGATACAGTCCCTTAGAA 1440
QY
1441 AGAGATAACTCAGAGAAATTTTTTAAACCGGAAGCAAGGCAAAACAGTTTACAGAGAA 1500
Db
1441 AGAGATAACTCAGAGAAATTTTTTAAACCGGAAGCAAGGCAAAACAGTTTACAGAGAA 1500
QY
1501 ATTGAGTCAAGTGCAGTACAAAGCTCGAGTGGCCCTGGAAATGATATGAGGTGAG 1560
Db
1501 ATTGAGTCAAGTGCAGTACAAAGCTCGAGTGGCCCTGGAAATGATATGAGGTGAG 1560
QY
1561 GAAGAAAAATACACAGAGTTTACAGAAATTCAGTGAACGTCAGGGGACACAGCATAAAC 1620
Db
1561 GAAGAAAAATACACAGAGTTTACAGAAATTCAGTGAACGTCAGGGGACACAGCATAAAC 1620
QY
1621 ACTAGGAAAAATAAATATATCTCTCTGGCAAGAAATAGAGAGTTCATATCTCTGGGA 1680
Db
1621 ACTAGGAAAAATAAATATATATCTCTCTGGCAAGAAATAGAGAGTTCATATCTCTGGGA 1680
QY
1681 AGTGGAGACAGAAATCACCGGTATGGCCAGCTCGATCGGCTCCATGCCATCAAGA 1740
Db
1681 AGTGGAGACAGAAATCACCGGTATGGCCAGCTCGATCGGCTCCATGCCATCAAGA 1740
QY
1741 TCACATTTCTACACTTCAGATTTCAACCCGAAATCTGGTTTCAGACCAAGAGTAGTTAAT 1800
Db
1741 TCACATTTCTACACTTCAGATTTCAACCCGAAATCTGGTTTCAGACCAAGAGTAGTTAAT 1800
QY
1801 GGAGGTGTTCCCTGGCCATCGCTTGGCCATCTCTCTCTCGCCACCTCTCTCGCTAC 1860
Db
1801 GGAGGTGTTCCCTGGCCATCGCTTGGCCATCTCTCTCTCGCCACCTCTCTCGCTAC 1860
QY
1861 CAGTCAGGTCCCAACTCTCTTCCACCTCGGGGACCAACCCCTACACGGCCGCCCTCCAGG 1920
Db
1861 CAGTCAGGTCCCAACTCTCTTCCACCTCGGGGACCAACCCCTACACGGCCGCCCTCCAGG 1920
QY
1921 CCCCCCTCGGGGACATCCAGACCCCGCTCTCAACCCCTCTGCTCATGTTCTCCAGTCTCT 1980
Db
1921 CCCCCCTCGGGGACATCCAGACCCCGCTCTCAACCCCTCTGCTCATGTTCTCCAGTCTCT 1980
QY
1981 GTCTCTACTATGCTTAACGATGCTTTCAGAGGGGCTCCAGGATGTCCTCCAAAGGCC 2040
Db
1981 GTCTCTACTATGCTTAACGATGCTTTCAGAGGGGCTCCAGGATGTCCTCCAAAGGCC 2040
QY
2041 CAGCGACATCTCGAAATACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100
Db
2041 CAGCGACATCTCGAAATACAGAGTTTCTGCTGGAGGGGTTCCATATCCAGTGGCCTA 2100
QY
2101 GAATTTGTATCCCAACACCCAGGTGAAGAGCTTACTCTCTCAGTGAAGAGCAAGT 2160
|||||

Db
2101 GAATTTGTATCCCAACCCACCACCCAGCTACTCTCTCAGTAGCAGGACCAGT 2160
QY
2161 CCTCTGGGGGAAACGTGTGTCATCAGTGTGTCAGTGGGTTCAGAGATTTATCCCTTAAACT 2220
Db
2161 CCTCTGGGGGAAACGTGTGTCATCAGTGTGTCAGTGGGTTCAGAGATTTATCCCTTAAACT 2220
QY
2221 CATAGACCAGGTCTCCAGACAGAAACAGTATTTGAAATACCCCCAGTGGGCGAGTCTT 2280
Db
2221 CATAGACCAGGTCTCCAGACAGAAACAGTATTTGAAATACCCCCAGTGGGCGAGTCTT 2280
QY
2281 GCTTCTCCCAAGCTGTATTTATTCAACTGAAGCTGTGGCATGCCATTTCCAGTGA 2340
Db
2281 GCTTCTCCCAAGCTGTATTTATTCCAACTGAAGCTGTGGCATGCCATTTCCAGTGA 2340
QY
2341 TCTCTAGCCCTGTAGTCTCGCATCGAACAGAGCTGTTACCCCTTCTAGTGAAGCTAAA 2400
Db
2341 TCTCTAGCCCTGTAGTCTCGCATCGAACAGAGCTGTTACCCCTTCTAGTGAAGCTAAA 2400
QY
2401 GATTTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAAGAAATATTTAA 2460
Db
2401 GATTTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCGAGGGAATAAAGAAATATTTAA 2460
QY
2461 CCCAATGAAACATACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCAACAGTTGT 2520
Db
2461 CCCAATGAAACATACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCAACAGTTGT 2520
QY
2521 TCTGAACATAGAAAAACAGATTTGATTTTAAAGAAATTTAAGAAATGATTTTAGTTTACAG 2580
Db
2521 TCTGAACATAGAAAAACAGATTTGATTTTAAAGAAATTTAAGAAATGATTTTAGTTTACAG 2580
QY
2581 CCAAGTTCTACTTCTGATCTATGATCAACTACTTAAACAAATAAGAGGAGAGAAAAA 2640
Db
2581 CCAAGTTCTACTTCTGATCTATGATCAACTACTTAAACAAATAAGAGGAGAGAAAAA 2640
QY
2641 TCAAGAGATTTGATCAAAAGACAAAAATGAAACAAAGTCTAAGGATTTCTTTTCAATTTGAAAAAT 2700
Db
2641 TCAAGAGATTTGATCAAAAGACAAAAATGAAACAAAGTCTAAGGATTTCTTTTCAATTTGAAAAAT 2700
QY
2701 AGCAGCAGCAACTGTACAGTGGCAGCAAGCCGAATAGCCCAAGCATTTCCCTTCA 2760
Db
2701 AGCAGCAGCAACTGTACAGTGGCAGCAAGCCGAATAGCCCAAGCATTTCCCTTCA 2760
QY
2761 ATACTTAGTAAACCGGACCAAGAGGGACCTGAGTCACTTCCCAAGGGTTTCAGACT 2820
Db
2761 ATACTTAGTAAACCGGACCAAGAGGGACCTGAGTCACTTCCCAAGGGTTTCAGACT 2820
QY
2821 TCCAGCCAGCATGTAAACAAAGAGAAAGACGATAGGAAAGAGAAAGACGAGCTGAG 2880
Db
2821 TCCAGCCAGCATGTAAACAAAGAGAAAGACGATAGGAAAGAGAAAGACGAGCTGAG 2880
QY
2881 CAAAGTTAGAAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCCAGTTCCTTCTCT 2940
Db
2881 CAAAGTTAGAAATCAACATTTGAATCCCAATGCAAGAGGTTCAACCCAGTTCCTTCTCT 2940
QY
2941 CAGCCAAAGCTTCTACTACCCCACTTCCCTCGGCTCAAGCAACCTAGCCCATCT 3000
Db
2941 CAGCCAAAGCTTCTACTACCCCACTTCCCTCGGCTCAAGCAACCTAGCCCATCT 3000
QY
3001 ATGGTGGGTCACTCAACAGCCAACTCCAGTTTATCTCAGCTGTTTGTGTTTGCACCAAAAT 3060
Db
3001 ATGGTGGGTCACTCAACAGCCAACTCCAGTTTATCTCAGCTGTTTGTGTTTGCACCAAAAT 3060
QY
3061 ATGATGTATCCAGTCCAGTGAGCCAGGGTGCAACCTTTATACCCAATACCTATGACG 3120
Db
3061 ATGATGTATCCAGTCCAGTGAGCCAGGGTGCAACCTTTATACCCAATACCTATGACG 3120
QY
3121 CCATGCGAGTGAATCAAGCAACATATAGAGAGTACCAAAATATGCCCGAACAGCGG 3180
Db
3121 CCATGCGAGTGAATCAAGCAACATATAGAGAGTACCAAAATATGCCCGAACAGCGG 3180
QY
3181 CAGAGCAGCATCATCAGAGTGCATGATGACCCAGCGTCAGCAGGGGGCCCAACCGATT 3240
Db
3181 CAGAGCAGCATCATCAGAGTGCATGATGACCCAGCGTCAGCAGGGGGCCCAACCGATT 3240

```
QY 3241 GCAGCCACCCACCCAGCTTACTCCAGGCAATATGTTGCTTACAGTCCCTCAGCAGTTCCCA 3300
Db 3241 GCAGCAACCCACCCAGCTTACTCCAGCAATATGTTGCTTACAGTCCCTCAGCAGTTCCCA 3300
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCACATTAATCAGTCTCAGCATCCCTCATCTATAGT 3360
Db 3301 AATCAGCCCTTGTTCAGCATGTGCCACATTAATCAGTCTCAGCATCCCTCATCTATAGT 3360
QY 3361 CTTGTAATACAGGGTAATGCTAGAAATGATGGCA CCAACACACAGCCCGCAGCCTGGTTTA 3420
Db 3361 CTTGTAATACAGGGTAATGCTAGAAATGATGGCA CCAACACACAGCCCGCAGCCTGGTTTA 3420
QY 3421 GTATCTTCTTCAGCAACTCAGTACGGGCTCATGAGCAGACGCATGCGATGATCATGT 3480
Db 3421 GTATCTTCTTCAGCAACTCAGTACGGGCTCATGAGCAGACGCATGCGATGATCATGT 3480
QY 3481 CCCAAATTACCATAACAACAGGAGACAAGCCCTTCTTCTTACTTTGGCCATTTCCACGGGC 3540
Db 3481 CCCAAATTACCATAACAACAGGAGACAAGCCCTTCTTCTTACTTTGGCCATTTCCACGGGC 3540
QY 3541 TCCCTTGTCTCAGCAGTATGCGCACCCCTAACGCTACCCCTGCACCCACATCTCCACACCT 3600
Db 3541 TCCCTTGTCTCAGCAGTATGCGCACCCCTAACGCTACCCCTGCACCCACATCTCCACACCT 3600
QY 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCCAAACATGGTGAAGTCACTCTGCA 3660
Db 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGCCAAACATGGTGAAGTCACTCTGCA 3660
QY 3661 CCCAGTCTCTTTCAGCAACATCAGACACAGGCGCGCCAGGCTCCCATCTGCGCAGTCCA 3720
Db 3661 CCCAGTCTCTTTCAGCAACATCAGACACAGGCGCGCCAGGCTCCCATCTGCGCAGTCCA 3720
QY 3721 CAGCAGCAGTCAAGCAATTTACCA CCGCGGGCTTGGCCAACTCCACCTCCATGACACT 3780
Db 3721 CAGCAGCAGTCAAGCAATTTACCA CCGCGGGCTTGGCCAACTCCACCTCCATGACACT 3780
QY 3781 GCCTCAACACGAGTGGCCACAGAAATGTTTCC CAGAGACACAACAGACTGTCTTTAG 3840
Db 3781 GCCTCAACACGAGTGGCCACAGAAATGTTTCC CAGAGACACAACAGACTGTCTTTAG 3840
QY 3841 ATCCATCTCTTCAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCACAGTCACT 3900
Db 3841 ATCCATCTCTTCAGTTCAGCGCGGTATACCAACCCACCCACATGCGCCACAGTCACT 3900
QY 3901 CAGGCTCATGTACAGTCAAGAAATGTTCTCTCATCCAACCTGCCCATCGCCCAATGATG 3960
Db 3901 CAGGCTCATGTACAGTCAAGAAATGTTCTCTCATCCAACCTGCCCATCGCCCAATGATG 3960
QY 3961 CTAATGACGACACAGCACCAGCGGTGCC CAGGCGCCCTGCTCAAAGTGCACTACAG 4020
Db 3961 CTAATGACGACACAGCACCAGCGGTGCC CAGGCGCCCTGCTCAAAGTGCACTACAG 4020
QY 4021 CCCATTCAGTCTCCACACAGCGCATTTCCCTATATAGCGCACCCCTTCAGTACAAGCC 4080
Db 4021 CCCATTCAGTCTCCACACAGCGCATTTCCCTATATAGCGCACCCCTTCAGTACAAGCC 4080
QY 4081 CACCAACCAACAGCAGTGTGAAGGCTGCGCTGGAGGAA CCGAAAGGCCAAATTCCTCTCTC 4140
Db 4081 CACCAACCAACAGCAGTGTGAAGGCTGCGCTGGAGGAA CCGAAAGGCCAAATTCCTCTCTC 4140
QY 4141 CTTTCTACTGCTTCTACCAACTGGGAAGCA CAGAAAACCTAGAAATTTCAATTTTGTGTTT 4200
Db 4141 CTTTCTACTGCTTCTACCAACTGGGAAGCA CAGAAAACCTAGAAATTTCAATTTTGTGTTT 4200
QY 4201 TAAATATATATGTTGATTTCTTGTAAACAT CCAATAGGAATGCTAACAGTTCACTTGAG 4260
Db 4201 TAAATATATATGTTGATTTCTTGTAAACAT CCAATAGGAATGCTAACAGTTCACTTGAG 4260
QY 4261 TGGAAAGATATTGGACCCAGTAGAGGCATTTTAGGAACTTTGGGGGTATTTCCATAATTCCA 4320
Db 4261 TGGAAAGATATTGGACCCAGTAGAGGCATTTTAGGAACTTTGGGGGTATTTCCATAATTCCA 4320
```

```
QY 4321 TATGCTGTTTTCAGAGTCCGCGAGGTACCC CAGCTCTGCTTCCCGAAACTGGAAGTTATTT 4380
Db 4321 TATGCTGTTTTCAGAGTCCGCGAGGTACCC CAGCTCTGCTTCCCGAAACTGGAAGTTATTT 4380
QY 4381 ATTTTAAATAAACCCCTTGAAAGTCAATGAACA CATCAGCTAGCAAAAGAGTAACAAGAGT 4440
Db 4381 ATTTTAAATAAACCCCTTGAAAGTCAATGAACA CATCAGCTAGCAAAAGAGTAACAAGAGT 4440
QY 4441 GATTCTTCTCTATTAATCTACTGCTAAAAA AAAAAA 4481
Db 4441 GATTCTTCTCTATTAATCTACTGCTAAAAA AAAAAA 4481
```

RESULT 4

```
US-10-802-228-1
; Sequence 1, Application US/10802228
; Publication No. US20050209178A1
; GENERAL INFORMATION:
; APPLICANT: Pulist, Stefan M
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF OBESITY
; FILE REFERENCE: 825466-100151
; CURRENT APPLICATION NUMBER: US/10/802,228
; CURRENT FILING DATE: 2004-03-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-802-228-1
```

```
Query Match 100.0%; Score 4481; DB 9; Length 4481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4481; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ACCCCGAGAAAGCAACCCAGCGCGCCGCTCTCACGTGTCCCTCCGCGGCCCGGG 60
Db 1 ACCCCGAGAAAGCAACCCAGCGCGCGCCGCTCTCACGTGTCCCTCCGCGGCCCGGG 60
QY 61 GCCACCTCACGTTCTGCTTCCGTTCTGACCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG 120
Db 61 GCCACCTCACGTTCTGCTTCCGTTCTGACCCCTCCGACTTCGGGTAAAGAGTCCCTATCCG 120
QY 121 CACCTTCGGTCCCAACCGCGGCCCTCGGCGCGCCGCCCTCCGATGCGTCAAGCGCCGCA 180
Db 121 CACCTTCGGTCCCAACCGCGGCCCTCGGCGCGCCGCCCTCCGATGCGTCAAGCGCCGCA 180
QY 181 GCTCCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTCGCGGAGCAGGTGGGCC 240
Db 181 GCTCCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTCGCGGAGCAGGTGGGCC 240
QY 241 GGGTGGCGCTCGCTCCAGCGCGCGCGGAGCGGCGGGCGGGCGGGTGGCGCGGCC 300
Db 241 GGGTGGCGCTCGCTCCAGCGCGCGCGGAGCGGCGGGCGGGCGGGTGGCGCGGCC 300
QY 301 CCGGACCGGTATCCCTTCGCGCGCCCTTCCCGCGCCCGGCCCGGCCCTCCCTCCCGG 360
Db 301 CCGGACCGGTATCCCTTCGCGCGCCCTTCCCGCGCCCGGCCCGGCCCTCCCTCCCGG 360
QY 361 CAGAGCTCGCTCCCTCGCGCTCAGACTGTTTGGTAGCAACGCGACGCGCGCGCG 420
Db 361 CAGAGCTCGCTCCCTCGCGCTCAGACTGTTTGGTAGCAACGCGACGCGCGCGCG 420
QY 421 TTTCCGGCCCGGCTCCCGCGGCTCTTGGTCTCGGCGGGCTCCCGGCCCTTCGTGCTC 480
Db 421 TTTCCGGCCCGGCTCCCGCGGCTCTTGGTCTCGGCGGGCTCCCGGCCCTTCGTGCTC 480
QY 481 GTCTTCTTCCCTTCGCGAGCCCGGGCGCCCTCCCGCGCGCCCAACCCGCGCTCCCGG 540
Db 481 GTCTTCTTCCCTTCGCGAGCCCGGGCGCCCTCCCGCGCGCCCAACCCGCGCTCCCGG 540
QY 541 CTCGCGCGCGGTGGTCCCGCGCGGTTCGGGCTCTCTTGGCGCGCGCGGCTCCCGGC 600
```


Db 541 CTCGGGGCCCGTGGTCCCGCGCGTTCGGCGGTCTCTTGGCGGCGCGGCTCCCGGC 600
Qy 601 TGTCCCGCCCGCGGTGGAGCGGTGTATGGGCCCCCTCACCATGTGCGTGAAGCCCCAG 660
Db 601 TGTCCCGCCCGCGGTGGAGCGGTGTATGGGCCCCCTCACCATGTGCGTGAAGCCCCAG 660
Qy 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720
Db 661 CAGCAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAG 720
Qy 721 CAGCGCGCGCGCGCTCCCAATGTCGCAAGCCCGCGGCGAGCGGCTTCTAGCGTGG 780
Db 721 CAGCGCGCGCGCGCGCTCCCAATGTCGCAAGCCCGCGGCGAGCGGCTTCTAGCGTGG 780
Qy 781 CCGCGCGCGCGCGCTTTCGGCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCG 840
Db 781 CCGCGCGCGCGCGCTTTCGGCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCGTCTCG 840
Qy 841 TCGTCTCGTCTCGCGGACCTCGCGCGGCGGAGCGCGGCTCGGCGAGAGGTGCAAAC 900
Db 841 TCGTCTCGTCTCGCGGACCTCGCGCGGCGGAGCGCGGCTCGGCGAGAGGTGCAAAC 900
Qy 901 AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAAAATAGGG 960
Db 901 AGTAACAAAGGACTGCTCAGTCTACGATTTCTTTTGTGGAATCTATGCAAAATAGGG 960
Qy 961 ATGGTTTCAATATCATCATCAGTTGTGGTCCAAATGTGAAGTCAAGTGAATAATGGA 1020
Db 961 ATGGTTTCAATATCATCATCAGTTGTGGTCCAAATGTGAAGTCAAGTGAATAATGGA 1020
Qy 1021 GGTATATATAGGAGGTTTTTAAACTTACAGTCCGAAGTGTGATTTGTGATCTGATGCC 1080
Db 1021 GGTATATATAGGAGGTTTTTAAACTTACAGTCCGAAGTGTGATTTGTGATCTGATGCC 1080
Qy 1081 GCACATGAGAAAGTACAGATCCAGTTCGGGCGGAAACGTGAAGAAATATGGAGGT 1140
Db 1081 GCACATGAGAAAGTACAGATCCAGTTCGGGCGGAAACGTGAAGAAATATGGAGGT 1140
Qy 1141 ATTTTCTTCAAATGTTTCAAGTCTTGTGTGTACAGTTTAAAGATATGCAATCCAGTTAT 1200
Db 1141 ATTTTCTTCAAATGTTTCAAGTCTTGTGTGTACAGTTTAAAGATATGCAATCCAGTTAT 1200
Qy 1201 GCAAAAGAGATCTTTTACTGACTCTGTATCAGTCTAAAGTGAATGGCGCAACAAA 1260
Db 1201 GCAAAAGAGATCTTTTACTGACTCTGTATCAGTCTAAAGTGAATGGCGCAACAAA 1260
Qy 1261 GAGAAAGACTTGGAGCCCTGGGATGAGGTGAATCAGCCCAATGAGGAACTTGAAGCT 1320
Db 1261 GAGAAAGACTTGGAGCCCTGGGATGAGGTGAATCAGCCCAATGAGGAACTTGAAGCT 1320
Qy 1321 TTGGAAATATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380
Db 1321 TTGGAAATATGACGTATCTAATGATGGGATCCCAATGATATGTTTCGATATATGAAGAA 1380
Qy 1381 AATTATGGTGTAGTGTCTAGTATGATAGCAGTTTATCTTTCGTATACAGTGCCTTAGAA 1440
Db 1381 AATTATGGTGTAGTGTCTAGTATGATAGCAGTTTATCTTTCGTATACAGTGCCTTAGAA 1440
Qy 1441 AGAGATAACTCAGAGAAATTTTAAACCGGGAAGCAGGCGCAACAGTTAGCAGAGAA 1500
Db 1441 AGAGATAACTCAGAGAAATTTTAAACCGGGAAGCAGGCGCAACAGTTAGCAGAGAA 1500
Qy 1501 ATTGAGTCAAGTGCACAGTACAAAGCTCGAGTGGCCCTGGAAATGATGATAGGATGAG 1560
Db 1501 ATTGAGTCAAGTGCACAGTACAAAGCTCGAGTGGCCCTGGAAATGATGATAGGATGAG 1560
Qy 1561 GAAGAAATATACAGCAGTTTCAAGAAATTTCCAGTGAACGTGAGGGGCAACAGATAAAC 1620
Db 1561 GAAGAAATATACAGCAGTTTCAAGAAATTTCCAGTGAACGTGAGGGGCAACAGATAAAC 1620
Qy 1621 ACTAGGAAATATATATATTTCTCTCGGCAAGAAATAGAGAGTCTATCTCTGGGA 1680
Db 1621 ACTAGGAAATATATATATTTCTCTCGGCAAGAAATAGAGAGTCTATCTCTGGGA 1680

Qy 1681 AGTGGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCGGCTCCATGCCATCAAGA 1740
Db 1681 AGTGGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCGGCTCCATGCCATCAAGA 1740
Qy 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAATTCCTGGTTCAAGACAAAGAGTAGTTAAT 1800
Db 1741 TCCACTTCTCACACTTCAGATTTCAACCCGAATTCCTGGTTCAAGACAAAGAGTAGTTAAT 1800
Qy 1801 GGAGGTGTTCCCTGGCCATCGCCTTGGCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
Db 1801 GGAGGTGTTCCCTGGCCATCGCCTTGGCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
Qy 1861 CAGTCAGTTCCT 1920
Db 1861 CAGTCAGTTCCT 1920
Qy 1921 CCCCCCTCGCGGCATCCAGACCCCCCTCTCACCCCTCTGCTCATGTTTCTCCAGTCTCT 1980
Db 1921 CCCCCCTCGCGGCATCCAGACCCCCCTCTCACCCCTCTGCTCATGTTTCTCCAGTCTCT 1980
Qy 1981 GTCTCTATCTATGCTCTAAAGCGCATGTTTTCAGAAAGGGCTTCCAAGGATGTCCCCAAAGGCC 2040
Db 1981 GTCTCTATCTATGCTCTAAAGCGCATGTTTTCAGAAAGGGCTTCCAAGGATGTCCCCAAAGGCC 2040
Qy 2041 CAGCGACATCTCTCGAAATCAGAGATTTCTGCTGGAGGGGTTTCCATATCCAGTGGCCTA 2100
Db 2041 CAGCGACATCTCTCGAAATCAGAGATTTCTGCTGGAGGGGTTTCCATATCCAGTGGCCTA 2100
Qy 2101 GAATTTGTATCCCAACAAACCCAGTCAGTGAAGCAGTACTCTCCAGTAGCAAGGACCACT 2160
Db 2101 GAATTTGTATCCCAACAAACCCAGTCAGTGAAGCAGTACTCTCCAGTAGCAAGGACCACT 2160
Qy 2161 CCCTCGGGGGAAACGTGGTTCATCAGTGGTTCAGTGGGTTTCCAAGATTTATCCCTTAAACT 2220
Db 2161 CCCTCGGGGGAAACGTGGTTCATCAGTGGTTCAGTGGGTTTCCAAGATTTATCCCTTAAACT 2220
Qy 2221 CATAGACCCAGGTCTCCAGACAGAACAGTATTTGAAATATCCCCAGTGGGCCAGTTCTT 2280
Db 2221 CATAGACCCAGGTCTCCAGACAGAACAGTATTTGAAATATCCCCAGTGGGCCAGTTCTT 2280
Qy 2281 GCTTCTCCCAAGCTGGTATTTATCCACTGAAGCTGTTGCCATGCTTATCCAGCTGCA 2340
Db 2281 GCTTCTCCCAAGCTGGTATTTATCCACTGAAGCTGTTGCCATGCTTATCCAGCTGCA 2340
Qy 2341 TCTCTCTACCCCTGTAGTCTCGATCGAACAGAGCTGTTTACCCCTTCTAGTGAAGCTTAA 2400
Db 2341 TCTCTCTACCCCTGTAGTCTCGATCGAACAGAGCTGTTTACCCCTTCTAGTGAAGCTTAA 2400
Qy 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCTGAGGGGAATTAAGAAATATTTAAA 2460
Db 2401 GATTCAGGCTTCAAGATCAGAGGAGAACTCTCTCTGAGGGGAATTAAGAAATATTTAAA 2460
Qy 2461 CCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAGTTGTT 2520
Db 2461 CCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAAGGTATATCACCAGTTGTT 2520
Qy 2521 TCTGAATATAGAAAAACAGATTTGATGATTTAAAGAAATTTAAGAAATGATTTTAGTTACAG 2580
Db 2521 TCTGAATATAGAAAAACAGATTTGATGATTTAAAGAAATTTAAGAAATGATTTTAGTTACAG 2580
Qy 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAAATAGAGAGGGAGAAAA 2640
Db 2581 CCAAGTTCTACTTCTGAATCTATGGATCAACTACTAAACAAAAATAGAGAGGGAGAAAA 2640
Qy 2641 TCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTGTCAAGGATTTCTTTTATTGAAAT 2700
Db 2641 TCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTGTCAAGGATTTCTTTTATTGAAAT 2700
Qy 2701 AGCAGCAGCAACTGTACCTAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTTCCCCTTCA 2760
Db 2701 AGCAGCAGCAACTGTACCTAGTGGCAGCAGCAAGCCGAATAGCCCCAGCATTTTCCCCTTCA 2760


```
QY 2761 ATACTTAGTAAACACGAGACCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACT 2820
Db 2761 ATACTTAGTAAACACGAGACCAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACT 2820
QY 2821 TCCAGCCCGAGCATGTAAACAGAGAGAAAGACGATAGGAGAGAGAGAAAGAGCGCAGCTGAG 2880
Db 2821 TCCAGCCCGAGCATGTAAACAGAGAGAAAGACGATAGGAGAGAGAGAAAGAGCGCAGCTGAG 2880
QY 2881 CAAGTTAGGAATCAACAAATTTGAATCCCAATGCAAAAGGAGTTCAACCCACGTTCTCTCT 2940
Db 2881 CAAGTTAGGAATCAACAAATTTGAATCCCAATGCAAAAGGAGTTCAACCCACGTTCTCTCT 2940
QY 2941 CAGCCAAAGCCCTTCTACTACCCCAACTTCACTCGGCCCTCAAGCAACAACCTTAGCCCACT 3000
Db 2941 CAGCCAAAGCCCTTCTACTACCCCAACTTCACTCGGCCCTCAAGCAACAACCTTAGCCCACT 3000
QY 3001 ATGGTGGGTCTCAACAGCAACTCCAGTTTATATCTCAGCCCTGTTTGTGTCACCAAAAT 3060
Db 3001 ATGGTGGGTCTCAACAGCAACTCCAGTTTATATCTCAGCCCTGTTTGTGTCACCAAAAT 3060
QY 3061 ATGATGTATCCAGTCCCAAGTCCAGCCGCGTGCACACCTTTATACCCAATACCTATGACG 3120
Db 3061 ATGATGTATCCAGTCCCAAGTCCAGCCGCGTGCACACCTTTATACCCAATACCTATGACG 3120
QY 3121 CCCATGCCAGTGAATCAAGCCCAAGCATATAGAGCAGTACCAATATGCCCCCAACAGCGG 3180
Db 3121 CCCATGCCAGTGAATCAAGCCCAAGCATATAGAGCAGTACCAATATGCCCCCAACAGCGG 3180
QY 3181 CAAGACCAAGCATCATCAGAGTGCATGATGACCCAGCGTCAAGCAGCGGCCCCACCGATT 3240
Db 3181 CAAGACCAAGCATCATCAGAGTGCATGATGACCCAGCGTCAAGCAGCGGCCCCACCGATT 3240
QY 3241 GCAGCAACCCCAACAGCTTACTCCACGCAATATGTTGCTACAGTCCCTCAGCAGTTCCCA 3300
Db 3241 GCAGCAACCCCAACAGCTTACTCCACGCAATATGTTGCTACAGTCCCTCAGCAGTTCCCA 3300
QY 3301 AATCAGCCCTTGTTCAGCATGTGCCATATTATCAGTCTCAGCATCTCCATGCTATAGT 3360
Db 3301 AATCAGCCCTTGTTCAGCATGTGCCATATTATCAGTCTCAGCATCTCCATGCTATAGT 3360
QY 3361 CCTGTAATACAGGGTAATGCTAGAAATGATGGCACCACCAACAGCCCGCCAGCCTGGTTTA 3420
Db 3361 CCTGTAATACAGGGTAATGCTAGAAATGATGGCACCACCAACAGCCCGCCAGCCTGGTTTA 3420
QY 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCGATGCGATGATGATGT 3480
Db 3421 GTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCGATGCGATGATGATGT 3480
QY 3481 CCCAAATTTACCATAACAAGAGACAAGCCCTTCTTCTACTTTGCCATTTTCCACGGGC 3540
Db 3481 CCCAAATTTACCATAACAAGAGACAAGCCCTTCTTCTACTTTGCCATTTTCCACGGGC 3540
QY 3541 TCCCTTGTCTCAGCATATGCGCACCTTAACGCTACCTGCAACCCACACATACCTCCACCCCT 3600
Db 3541 TCCCTTGTCTCAGCATATGCGCACCTTAACGCTACCTGCAACCCACACATACCTCCACCCCT 3600
QY 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGAGCCAAACATGTTGGAAGTCATCCTGCA 3660
Db 3601 CAGCCTTCAGCTACCCCACTGGACAGCAGCAAGAGCCAAACATGTTGGAAGTCATCCTGCA 3660
QY 3661 CCCAGTCTGTTCAGCACCATCAGCACCGCGGCCAGGCTCTCCATCTCGGCCAGTCCA 3720
Db 3661 CCCAGTCTGTTCAGCACCATCAGCACCGCGGCCAGGCTCTCCATCTCGGCCAGTCCA 3720
QY 3721 CAGCAGCAGTCAGCCATTTACACGCGGGCTTGCGCCAACTCCACCCCTCCATGACACCT 3780
Db 3721 CAGCAGCAGTCAGCCATTTACACGCGGGCTTGCGCCAACTCCACCCCTCCATGACACCT 3780
QY 3781 GCCTCCAAACAGCGAGTGCCTCCAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840
Db 3781 GCCTCCAAACAGCGAGTGCCTCCAGAAATAGTTTCCAGCAGCAGCAACAGACTGTCTTTACG 3840
QY 3841 ATCCATCTCTTCAGGTTACGCCGGCGTATACCAACCCACCCACATGGGCCACGTACCT 3900
```

```
Db 3841 ATCCATCTCTTCAGGTTACGCCGGCGTATACCAACCCACCCACATGGGCCACGTACCT 3900
QY 3901 CAGGCTCATGTACAGTCAGGAATGTTCTCTTCTCATCCAACTGCGCCATGCGCCCAATGATG 3960
Db 3901 CAGGCTCATGTACAGTCAGGAATGTTCTCTTCTCATCCAACTGCGCCATGCGCCCAATGATG 3960
QY 3961 CTAATGACGACACAGCCACCCGGCGGTCCCGAGCGCCCTCGCTCAAAGTGCACTACAG 4020
Db 3961 CTAATGACGACACAGCCACCCGGCGGTCCCGAGCGCCCTCGCTCAAAGTGCACTACAG 4020
QY 4021 CCCAATTCAGTCTCGAACAACAGCCGATTTCCCTCTATATGACGACCCCTTCAGTACAGCC 4080
Db 4021 CCCAATTCAGTCTCGAACAACAGCCGATTTCCCTCTATATGACGACCCCTTCAGTACAGCC 4080
QY 4081 CACCACCAACAGCAGTGTGTAAGGCTGCCCTGAGGAAACCGAAAGGCCAAATTCCTCTCTC 4140
Db 4081 CACCACCAACAGCAGTGTGTAAGGCTGCCCTGAGGAAACCGAAAGGCCAAATTCCTCTCTC 4140
QY 4141 CTTTCTACTGTTCTTACCAACTGGAAGCACAGAAAACTAGAAATTTCAATTTATTTGTTTT 4200
Db 4141 CTTTCTACTGTTCTTACCAACTGGAAGCACAGAAAACTAGAAATTTCAATTTATTTGTTTT 4200
QY 4201 TAAATATATATGTTGATTTCTTCTTAACATCAATAGGAATGCTTAACAGTTTCACTTGCAG 4260
Db 4201 TAAATATATATGTTGATTTCTTCTTAACATCAATAGGAATGCTTAACAGTTTCACTTGCAG 4260
QY 4261 TGGAGATATCTTGGACCGAGTAGAGGCAATTTAGGAACCTTGGGGCTATTCCATAAATCCA 4320
Db 4261 TGGAGATATCTTGGACCGAGTAGAGGCAATTTAGGAACCTTGGGGCTATTCCATAAATCCA 4320
QY 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGTCTGCTTCCGAAACTGGAAGTTATTT 4380
Db 4321 TATGCTGTTTCAGAGTCCCGCAGGTACCCAGTCTGCTTCCGAAACTGGAAGTTATTT 4380
QY 4381 ATTTTAAATACCCCTTGAAGTCAATGAACATCATGAACATCAGTAGCAAAAGAGTAACAAGAGT 4440
Db 4381 ATTTTAAATACCCCTTGAAGTCAATGAACATCATGAACATCAGTAGCAAAAGAGTAACAAGAGT 4440
QY 4441 GATTCTTGTCTCTATTACTGCTAAAAAATAAAAAAAAAAAAAA 4481
Db 4441 GATTCTTGTCTCTATTACTGCTAAAAAATAAAAAAAAAAAAAA 4481

RESULT 5
US-10-141-541-3
; Sequence 3, Application US/10141541
; Publication No. US20030167495A1
; GENERAL INFORMATION:
; APPLICANT: Pulat, Stefan M.
; TITLE OF INVENTION: SCA2 Knockout Animal and Methods of Use
; FILE REFERENCE: P-CE 5244
; CURRENT APPLICATION NUMBER: US/10/141,541
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/289,231
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4225
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (27) ... (3884)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (222) ... (224)
; OTHER INFORMATION: ccs=Pro
US-10-141-541-3
```

Query Match 70.0%; Score 3138.2; DB 6; Length 4225;
Best Local Similarity 86.2%; Pred. No. 0;

Matches 3662; Conservative 1; Mismatches 473; Indels 112; Gaps 14;		
QY	222	CGCCGAGCCAGGTGGCCCGGGTGGCGCTCGCTCCAGCGGCCCGGCGCGCGAGCGGGCG 281
DB	71	CCCGAGCGCGCGCGCGCGGGCTGGGCGCGCGGGCGCTCGCTCCCGCGAGCGCGCG 130
QY	282	GGCGGGCGGTGGCGGGCCCGGGACCGGTATCCCTCCGCGCGCCCTCCCGCCCGCGGGCC 341
DB	131	GGCGGGCGGGCGGGCGGGTGGCGGTATCCCTCCGCGCGCCCTCCCGCCCGCGGGGCC 190
QY	342	CGGCCCCCTCCCTCCCGGACAGCTCGCTCCCTCCGCTCAGACTGTTTTGGTAGCAA 401
DB	191	CGGGGCCCTCCCGGGGCGGGCTCGGCAAGSTGGGCTCAGACTGTTTTGGTAGCAA 250
QY	402	CGGCAACGCGGGCGGGCGGTTCCTGGGCCCGGCTCCCGGGCGCTCCTTGGTCTCGGGGGCC 461
DB	251	CGGGCA-----CGGGCGGTCCCGGCCCGGCTCCCGGGCGGCTCGTGGTGTCTGGGGGCC 304
QY	462	TCCCGCCCTTCGTCTGTCGTCTCTCCCTCCGCTCCGCGAGCCCGGGGGCCCTCCCGGGCCG 521
DB	305	TCCCGCCCTTCGTCTGTCGTCTCTCTGCTCTG-----GCCCGGGCGGGCCAC 352
QY	522	GCAACCGCGGCTCCCGCTCGGCGCCCGTGGCTCCCGCGCGCTTCGGGGCTCTCTT 581
DB	353	GCGGGCCCGGGCTGCGCGCGCGGGCTCGCGGGTCCCGCGCGCTCCCGGGCTCTCTCTC 412
QY	582	---GGCGCGCCCGGCTCCCGGGTGTCCCGCGCGGTGCGAGCCCGGTGTATGGGGCCCT 638
DB	413	CTCGGCGCGCGGCAACCGGGTGTCCCGCGCGGGTGGAGCCGGTGTATGGGGCGCT 472
QY	639	CACCATGTCTGTAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 698
DB	473	CACCATGTCTGTAAGCC-----490
QY	699	ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 758
DB	491	-----GAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 532
QY	759	CGGCGCGGCTCTAGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTC 818
DB	533	CGGCGGCTCTCTCTCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCTC 592
QY	819	CTCGTCTCGGCCACCGGCTCCCTCTCTCGTG-----GTGCGGGCGACCTCGGGCGGGGAG 875
DB	593	CGTGGTTCGCGGCCCGCGCGCGCGTGGCGTCTCTCTCGCGCGCGCGCGCGCGCGCG 652
QY	876	GCCCGGCTGGGCGAGAGTGAACAGTAACAAAGGACTGCTCAGTCTACGATTTCTTT 935
DB	653	TCCCGGCTGGGCGAGAGTGGAAACAGTAGCAAGGACTGCTCAGCCTACGATTTCTTT 712
QY	936	TGATGGAATCTATGCAATATAGGATGGTTTATATATATATATATATATATATATATAT 995
DB	713	TGATGGAATCTATGCAACGTGAGGATGGTTTATATATATATATATATATATATATAT 772
QY	996	ATGTGAAGTACAAGTGAAGTATATATGAAGGAGTTTTTAAACTTACAGTCC 1055
DB	773	ATGTGAAGTACAAGTGAAGTATATATGAAGGAGTTTTTAAACATACAGTCC 832
QY	1056	GAAATGTGATTTGGTACTTGCATCGCGCATATGAGAAAGTACAGAAATCCAGTTTCGGGGCC 1115
DB	833	TAAATGTGATTTGGTACTTGCATCGCGCATATGAGAAAGTACAGAAATCCAGTTTCGGGGCC 892
QY	1116	GAAACGTGAAGAAATATGAGAGTATTTTGTTCAAATGTTTCAGACTTGTGTGGTGA 1175
DB	893	AAGCGTGAAGAAATATGAGAGTATTTTGTTCAAATGTTTCAGACTTGTGTGGTGA 952
QY	1176	GTTTAAAGATATGAGTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAG 1235
DB	953	GTTTAAAGATATGAGTCCAGTTATGCAAGAGAGATGCTTTTACTGACTCTGCTCTCAG 1012
QY	1236	TGCTAAAGTGAATGGCGAAACAAAGAGAGAGCCTGGAGCCCTGGGATGCGAGTGA 1295
DB	1013	CGCAAGGTTGAATGGTGGAGCAGAGGAGAGGAGCCTGGAGCCCTGGGATGCGAGGAGCT 1072

QY	1296	CACAGCCATGAGGAACCTTGAGGCTTTGGAATAACGCTATCTAATGATGGATGCCAAC 1355
DB	1073	CACGGCCAGCGAGGAGCTGGAG---CTGGAGAAATGATGTGTCTAATGGATGGACCCCAA 1129
QY	1356	TGATATGTTTCGATATAATGAAGAAATATATGGTGTAGTGTCTAGCTATGATAGCAGTTT 1415
DB	1130	TGACATGTTTCGATATAATGAAGAAATATATGGTGTAGTGTCTACATATGATAGCAGTTT 1189
QY	1416	ATCTTCGTATACAGTGCCTTTAGAAAGAGATAACTCAGAAGAAATTTTAAAAACGGGAAGC 1475
DB	1190	ATCTTCATATACGTTTCCCTTTAGAAAGGACAACTCAGAAGAAATTTCTTTAAACGGGAGC 1249
QY	1476	AAGGCAAAACAGTTAGCAGAAAGAAATTTAGTCAAGTGCAGTCAAAAGCTCGATGCG 1535
DB	1250	AAGGCAAAACAGTTAGCAGAAAGAAATTTGAAATCCAGTGTCTAGTCAAAAGCTCTGTGCG 1309
QY	1536	CCTGGAATAATGATAGGAGTGAGAGGAGAAATAACACAGCAGTTTCAGAGAAATTTCCAG 1595
DB	1310	CCTTGAGAAATGATACCGAGTGAGAGGAGAAATAACACAGCAGTTTCAGAGAAATTTCCAG 1369
QY	1596	TGAACGTGAGGGGACAGCATAAACACATAGGGAAATATAATATATATTTCTCTGGACAAAG 1655
DB	1370	TGACCGGAGGGGCGATGGCCCCAACACTAGGGACAAATAATATATTTCTCTCTGGACAAAG 1429
QY	1656	AAATAGAGAAATCATATCTCTGGGGAAGTGGGAGACAGAAATTCACCGCTATGGGCCAGCC 1715
DB	1430	AAACAGAGAAATCTATCTCTGGGGAAGTGGGAGACAGAGCTCACCAAGGATGGGCCAGCC 1489
QY	1716	TGGATCGGGCTCCATGCGCATCAAGATCCACTTTCACACTTCAGATTTTCAACCCGAAATTC 1775
DB	1490	TGGGCGAGGCTCCATGCGCTCAAGAGCTGCTTCTTACACTTCAGATTTTCAACCCGAAATTC 1549
QY	1776	TGGTTTCAGAACAAAGAGTAGTTAAATGGAGGTGTTCCTCGCCCATCGCTTGCCCATCTCC 1835
DB	1550	TGGCTCAGAACAAAGAGTAGTTAAATGGAGGTGTTCCTCGCCCATCGCTTGCCCATCTCA 1609
QY	1836	TTCTCTCGGCCACCTTCTCGCTACAGTCAAGTTCGCAACTCTCTTCCAGCTCGGGGAGC 1895
DB	1610	TTCTCTCGGCCACCTTCTCGCTACAGTCAAGTTCGCAACTCTCTTCCAGCTCGGGGAGC 1669
QY	1896	CACCCCTACACGGCGCGCTCCAGGCCCGCTCGCGGCCATCCAGACCCCGCTCTCACCC 1955
DB	1670	CACCCCTACACGGCGCGCTCCAGGCCCGCTCGAGGCCATCCAGACCCCGCTCTCACCC 1729
QY	1956	CTCTGCTCATGTTTCTTCAGCTCTCTCTACTATGCTTAAACGCAATGCTTTCAGAAAG 2015
DB	1730	CTCTGCTCATGTTTCTTCAGCTCTCTCTACTATGCTTAAACGCAATGCTTTCAGAAAG 1789
QY	2016	GCCTCAAGGATGTCCCAAGGCCCGAGGACATCTCGAAATCACAGAGTTTCTGCTGG 2075
DB	1790	ACCCCAAGGATGTCTCCAAAGGSCACAGCGCCACCTCGGAATCACAGAGTCTCTGCTGG 1849
QY	2076	GAGGGTTTCATATCCAGTGGGCTAGAAATTTGTATCCCAACCCACCCAGTCAAGCAGC 2135
DB	1850	GAGAGGCTTCATGTCTAGTGGGCTAGAAATTTGTATCCCAATATCCCGCAAGTGAAGCAGC 1909
QY	2136	TACTCTCTCAGTAGCAAGGACAGTCCCTCGGGGGGAAAGTGTGTCTCAGTGGTCAAGTGG 2195
DB	1910	TGCTCTCTCAGTAGCAAGGACAGTCTCTGAGGGGAAAGTGTGTCTCAGTGGTCAAGTGG 1969
QY	2196	GGTTCCAAAGTATATCCCTTAAACTCATAGACCCAGGTCTCCAGACAGAACAGTATGG 2255
DB	1970	GGTTCCAAAGTATATCTCCCAAAATCACAGACCCAGGTCTCCAGGCGAGAGCAGATTGG 2029
QY	2256	AAATACCCCGAGTGGGCGAGTTCTTGGTCTCCCAAGCTGGTATTTATTCCAACTGAAC 2315
DB	2030	AAACTCTCCAGCGGGGCTGTGCTTCTCCCAAGCTGGGATCATCTCCCTGAGAAAGC 2089
QY	2316	TGTTGGCATGCTTCTCCAGCTGCATCTCTCTAGCTGTCTGCTAGTCTCGATCGAACAGC 2375
DB	2090	CGTTTCCATGCTGTTCCTCCCGCCGCACTCTCGACTCTCTGCGAGCCCTGCAATCCAAACAGC 2149

QY	2376	TGTTACCCCTTCTAGTGAGGCTAAAGATTTCCAGCGTTTCAAGATCAGAGGCGAAGACTCTCC	2435
DB	2150	ACTGACCCCATCTTATTGAGGCAAAAGATTTCCAGCGTTTCAAGATCAGAGGCGAAGACTCTCC	2209
QY	2436	TGCAGGGAAATAAGAAATAATTAACACCAATGAACACATCACTAGCTTCTCAAAAGCTGA	2495
DB	2210	TGCGAGGAGTAAAGAAATGTTAAAGCAAGTGAACATCACCTAGCTTTTCAAAGCTGA	2269
QY	2496	AAACAAAGGTATATCACACAGTTGTTTCTTGAAACATAGAAAAAGATTTGATTTAAAGAA	2555
DB	2270	CAACAAAGGTATGTCACACAGTTGTTTCTGGAACAAGAAAAAGATTTGATTTAAAGAA	2329
QY	2556	ATTTAAAGAAATGATTTTAGTTTAGTACAGCCAAAGTCTACTTCTGAATCTATGATCAACTACT	2615
DB	2330	GTTTAAAGAAATGATTTTAGTTTAGTACAGCCAAAGCTCTACATCTGAATCTATGATCAACTACT	2389
QY	2616	AAACAAAAATAGAGAGGAGAAAAAATCAAGAGATTTGATCAAGACAAAAATTTGAACCAAG	2675
DB	2390	AAGCAAAATAGAGAGAGAAAAAGTACAGAGATTTGATTAAGATTAATAACGGAAGCAAG	2449
QY	2676	TGCTAAGGATTTCTTTTCATTTGAAAAAT-----AGCAGCAGCAACTGTACCTAGCTGGCAGCAG	2729
DB	2450	TGCTAAGGATGATTTTCATTGACAGCAGCAGCAGCAGCAGCBACTGTATACAGTGGCAGCAG	2509
QY	2730	CAAGCCGAATAGCCCCAGCATTTCCCTTTCAATACTTAGTAAACAAGGACACAAAGAGGGG	2789
DB	2510	CAAGACCAACAGCCCTAGCATCTCCCTTCCATGCTTAGTAATGACAGACACAAGAGGGG	2569
QY	2790	ACCTGAGGTCACTTTCCCAAGGGGTTCCAGACTTCCAGCCCAGCATGTAAACAAGNAGAAGA	2849
DB	2570	GCCTGAGGTCACTTCCCAAGGGGTCAGACTTCCAGCCCAGCCTGCAACAAGAGNAGAAGA	2629
QY	2850	CGATAAGGAAGAGAAGACGCGCTGAGCAAGTTAGGAAATCAACATTGAAATCCCAA	2909
DB	2630	TGACAGAGAGAGAAAGAAACAACAAGCAGCGTTAGGAAATCGAAATTCGAATTCCAA	2689
QY	2910	TGCAAAAGGAGTTCAACCCACGTTCTTCTTCAGCCAAAGCCTTCTACTACCCCAACTTC	2969
DB	2690	TGCAAAAGGAGTTCAACCCCTCGTTCTTTCTTCAGCCAAAGCCTTCTACTACCCCAACGTC	2749
QY	2970	ACCTCGGCCTCAAGCACAACTAGGCCATCTATGTTGGTGGTTCATCAACAGCGAACTCCAGT	3029
DB	2750	ACCTCGGCCTCAAGCACAAACCCAGCCCATCTATGTTGGGTTCATCAGCAGCGCAGCTCCAGT	2809
QY	3030	TTATACTCAGCGCTGTTGTTTGACCAAAATATCATGTATCCAGTCCAGTCCAGTGAGCCCCAGG	3089
DB	2810	GTAACACTCAGCGCTGTGTGCTTCGACCCCAATATGATATCCCGTCCAGTGAGCCCCGGG	2869
QY	3090	CGTGCAACCTTTATACCAATACCTATGAGCCCATGCGGTGAATCAAGCCCAAGACATA	3149
DB	2870	CGTACAACCTTTATACCAATACCTATGAGCCCCATGCTGTGAACCAAGCCCAAGACATA	2929
QY	3150	TAGAGC-----AGTACCAATATGCCCAACAGCGGCAAGACAGCATATCAGAGTGC	3203
DB	2930	TAGAGCAGGTAAAGTACCAAAATATGCCCAACAGCGCAACAGCCAAACATCATCAAAAGCAC	2989
QY	3204	CATGATGACCCAGCGTCAGCAGCGGGCCACCGATTGAGCCACCCCAACCGCTTACTC	3263
DB	2990	CATGATGACCCAGCGTCCGCGGAGGGCCACCCATCGTAGCCACCCCGCCGCTTACTC	3049
QY	3264	CACGCAATATGTTCCTACAGTCTCTCAGCAGTTCCCAAAATCAGCGCCCTTGTTCAGCATGT	3323
DB	3050	CACCTCAGTACGTTGCCTACAGCCCTCAGCAGTTTCCCAATCAGCGCTTGTTCAGCATGT	3109
QY	3324	GCACAATTATCAGTCTCAGCATTCCTCATGTCTATAGTCCGTGAATACAGGGTAATGCTAG	3383
DB	3110	GCCGCAATTATCAGTCTCAGCATTCCTCATGTGTACAGTCTCTGTACATCAAGGTAAATGCCAG	3169
QY	3384	AATGATGCAACCAACACAGCCCGCAGCTGGTTAGTATCTCTTCAGCACTCAGTA	3443
DB	3170	GATGATGGCAACCAAGCACATGTCTCAGCCTGGTTTATGTGTCTTCTTCAGCTGCTCAGTT	3229
QY	3444	CGGGGCTCATGAGCAGACGCGATGGATGTATGCAATGTCACCAATTAACATACCAACAGGA	3503

Db	3230	CGGGGCTCAGAGCAGACGCACGTATGATGTCCCAAAATTACATACACAAGGA	3289
QY	3504	GACAAGCCCTTCTTCTACTTTGCGATTTCCACGGGTCCTCTGCTCAGCAGTATGCGCA	3563
Db	3290	GACAAGCCCTTCTTCTACTTTGCGATTTCCACGGGTCCTCTGCTCAGCAGTATGCGCA	3349
QY	3564	CCCTAAAGCTACCTGTGACCCACATCTCCACACCTCTAGCCTTCAGTACCCTCCACTGG	3623
Db	3350	TCCTTAATGCGGCGCTGATCCACATATCCCATCTCTCAGCCTTCGCGCACTCCCA	3409
QY	3624	ACAGCAGAACCCAAATGCTGGAAGTCACTCTGCACCCAGTCTGTTTCAGACACCATCA	3683
Db	3410	ACAGCAGAACCCAGCATGCTGGAAGTCACTCTGCACCCAGTCTGTTTCAGACACCATCA	3469
QY	3684	GCACAGGCGCCAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTCAAGCCATTACCA	3743
Db	3470	GCACAGGCTGCCAGGCTCTTCTATCTGGCCAGTCCACAGCAGCAGTTCGCCATTATCA	3529
QY	3744	CGCGGGCTTGGCCAACTCCACCTCCATGACACGCTGCTCCAAACGCGAGTCCGCACA	3803
Db	3530	TGCGGGGCTGGCAACAACACCTTCCATGACACCTGCTCTAAACACAGTCTCCACA	3589
QY	3804	GAATAGTTTCCCAGCAGCAACACAGACTGCTTTTACGATCCATCTCTCACGTTTCAGCC	3863
Db	3590	GAGCAGTTTCCCAGCAGCAACACAGCAGTCTTCCACATCCACCTTCTCATGTTTCAGCC	3649
QY	3864	GGCGTATACCAACCCACCCACATGGGCCAGTACCTCAGGCTCATGTACAGTCAGGAAT	3923
Db	3650	GGCATACACACCCACCCACATGGGCCAGTACCTCAGGCTCATGTACAGTCAGGAAT	3709
QY	3924	GGTTCTTCTCATCCAACTGCCATCGCCAAATGATGCTAATCAGCAGCAGCCACCCGG	3983
Db	3710	GGTTCTTCTCATCCAACTGCCATCGCCAAATGATGCTAATCAGCAGCAGCCACCC	3766
QY	3984	CGGTCCCAGGCGCCCTCGCTCAAAAGTGCACTACAGCCCATTCAGTCTCGACAAACGC	4043
Db	3767	CGGTCCCAGGCGCCCTCGCTCAAAAGTGCACTACAGCCCATTCAGTCTCGACAAACGC	3826
QY	4044	GCATTTCCCTTATATGACGACCTTTCAGTACAAGCCACACCAACAGCAGTTGTAAGG	4103
Db	3827	GCATTTCCCTTATATGACGACCTTTCAGTACAAGCCACACCAACAGCAGTTGTAAGG	3886
QY	4104	CTGCCCTGGAGAACCGAAAGGCCAAATCCCTCCCTCCCTTCTACTGTTCTACCAACTG	4163
Db	3887	CTGCCCTGGAGAACCGAAAGGCCAAATCCCTCCCTCCCTTCTACTGTTCTACCAACTG	3945
QY	4164	GAAGCAGAGAACTAGAAATTTCATTTATTTGTTTTTAAATATATATGTCATTTCTT	4223
Db	3946	GAAGCAGAGAACTAGAACTTCATTTATTTGTTTTTAAATATATATGTCATTTCTT	3999
QY	4224	GTAACATCCAAATAGGAATGCTTAACAGTTCACCTTGCAGTGGGAAGAT-ACTTGGACCCAGTA	4282
Db	4000	TTAACACTCATAGGAATGCTTAACAGTTCACCTTGCAGTGGGAAGATGTTTTGGACCCAGTA	4059
QY	4283	GAGGCATTTAGGAACCTTGGGGCTATTCCATAATTCCATATGCTGTTTCAGAGTCCGCA	4342
Db	4060	GAGGCATGTAAGGCACTT-GTGGGCTGTTCCATAAATCCATGCTGTTTGCAGGGTCTTGCA	4118
QY	4343	GGTACCCCGACTCTGCTTGCAGAACTGGAAGTATTATTTTAAATACCCCTTGAAG	4402
Db	4119	AGTA-CCGACTCTGCTGCTGAACTGGAAGTATTATTTTAAATGGCCCTTGAGAG	4177
QY	4403	TCATGAACACATCAGCTAGCAAAAGAAAGTAAACAGAGTGATTCCTGCT	4450
Db	4178	TCATGAACACATCAGCTAGCAACAGAAAGTAAACAGAGTGATTCCTGCT	4225

RESULT 6
US-09-925-298-15
; Sequence 15, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:

DB	1746	ATTCTTGCTGCTATTACTGCTTA	AAAAAAAAAAAAAAAAAAAAA	1785
RESULT 7				
US-10-102-806-15				
; Sequence 15, Application US/10102806				
; Publication No. US20030054421A1				
; GENERAL INFORMATION:				
; APPLICANT: Rosen et al.				
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies				
; FILE REFERENCE: PA103PIC1				
; CURRENT APPLICATION NUMBER: US/10/102,806				
; CURRENT FILING DATE: 2002-03-22				
; PRIOR APPLICATION NUMBER: 09/4925,298				
; PRIOR FILING DATE: 2001-08-10				
; PRIOR APPLICATION NUMBER: PCT/US00/05881				
; PRIOR FILING DATE: 2000-03-08				
; PRIOR APPLICATION NUMBER: 60/124,270				
; PRIOR FILING DATE: 1999-03-12				
; NUMBER OF SEQ ID NOS: 846				
; SOFTWARE: PatentIn Ver. 2.0				
; SEQ ID NO 15				
; LENGTH: 2006				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURES:				
; NAME/KEY: misc feature				
; LOCATION: (862)				
; OTHER INFORMATION: n equals a,t,g, or c				
; NAME/KEY: misc feature				
; LOCATION: (1006)				
; OTHER INFORMATION: n equals a,t,g, or c				
US-10-102-806-15				
Query Match				
Best Local Similarity 38.6%; Score 1730.6; DB 5; Length 2006;				
Matches 1768; Conservative 2; Mismatches 8; Gaps 3;				
QY	2710	AACGTACAGTGGCAGCAAGCGGAATAGCCCGCAAGTTC	CCCCCAGCAATTTCC	CCCTTTCAATACTTAGT 2769
DB	6	AGCTGTACCACTGGCAGCAGCAAGCGGAATAGCCCGCAAGTTC	CCCCCAGCAATTTCC	CCCTTTCAATACTTAGT 65
QY	2770	AACACGGAGCAACAGAGGGGACCTGAGGTCACTTCCCAAGGGGTT	CAGACTTCCAGCCCA	2829
DB	66	AACACGGAGCAACAGAGGGGACCTGAGGTCACTTCCCAAGGGGTT	CAGACTTCCAGCCCA	125
QY	2830	GCATCTAACAGAAAGACGATGAAGAGAGAGAAAGACGACGCTGACCAAGTTAGG		2889
DB	126	GCATCTAACAGAAAGACGATGAAGAGAGAGAAAGACGACGCTGACCAAGTTAGG		185
QY	2890	AAATCAACATTTGAATCCCAATGCAAGGAGTTCAACCCACAGTTCCCTCTCAGCCAAAG		2949
DB	186	AAATCAACATTTGAATCCCAATGCAAGGAGTTCAACCCACAGTTCCCTCTCAGCCAAAG		245
QY	2950	CCTTCTATACCCCAACTTCACTCGGCTCAAGCAACCTAGCCCACTATATGTTGGGT		3009
DB	246	CCTTCTATACCCCAACTTCACTCGGCTCAAGCAACCTAGCCCACTATATGTTGGGT		305
QY	3010	CATCAACGCCAACTCCAGTTTATATCTAGCCTGTTTGTGTCACCAAAATATGATGTAT		3069
DB	306	CATCAACGCCAACTCCAGTTTATATCTAGCCTGTTTGTGTCACCAAAATATGATGTAT		365
QY	3070	CCAGTCCCAAGTGAGCCCAAGGGGTGCAACTTTATACCCCAATACCTATGACGCCCAATGCCA		3129
DB	366	CCAGTCCCAAGTGAGCCCAAGGGGTGCAACTTTATACCCCAATACCTATGACGCCCAATGCCA		425
QY	3130	GTGAATCAAGCCCAAGACATATAGAGC-----AGTACCAAAATATGCCCCCAACAGCGGCA		3183
DB	426	GTGAATCAAGCCCAAGACATATAGAGC-----AGTACCAAAATATGCCCCCAACAGCGGCA		485
QY	3184	GACAGCATCATCAGGTGCCATGATGACCCAGCGTTCAGCAGGGGCCACCCGATTTGCA		3243
DB	486	GACAGCATCATCAGGTGCCATGATGACCCAGCGTTCAGCAGGGGCCACCCGATTTGCA		545

QY	3244	GCCACCCACACAGCTTACTCCACGCAATATGTTGCTTACAGTCTCTCAGCAGTTCCCAAT	3303
DB	546	GMACCCACACAGCTTACTCCACGCAATATGTTGCTTACAGTCTCTCAGCAGTTCCCAAT	605
QY	3304	CAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCTCATGTCTATAGTCTCT	3363
DB	606	CAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAGCATCTCATGTCTATAGTCTCT	665
QY	3364	GTAATACAGGGTAATGCTAGAAATGATGCGCACCAACACACAGCCCGAGCCTGGTTTAGTA	3423
DB	666	GTAATACAGGGTAATGCTAGAAATGATGCGCACCAACACACAGCCCGAGCCTGGTTTAGTA	725
QY	3424	TCCTTTTTCAGCAACTCAGTAGCGGGCTCATGAGCAGACGATGCGATGATGATGTCCTCC	3483
DB	726	TCCTTTTTCAGCAACTCAGTAGCGGGCTCATGAGCAGACGATGCGATGATGATGTCCTCC	785
QY	3484	AAATTAACATACAAAGAGGACAAAGCCCTTCTTTTCTACTTTTGGCATTTTCCACGGCTCC	3543
DB	786	AAATTAACATACAAAGAGGACAAAGCCCTTCTTTTCTACTTTTGGCATTTTCCACGGCTCC	845
QY	3544	CTTGCTCAGCAGTATGCGC-ACCCTAAGCTACCTGCACCCACCATACTCCACACCTCA	3602
DB	846	CTTGCTCAGCAGTATGCGC-ACCCTAAGCTACCTGCACCCACCATACTCCACACCTCA	905
QY	3603	GCCTTCAGCTACCCCTCAGTGGACAGCAAGCAACATGGTGGAAAGTCACTCTGCAACC	3662
DB	906	GCCTTCAGCTACCCCTCAGTGGACAGCAAGCAACATGGTGGAAAGTCACTCTGCAACC	965
QY	3663	CAGTCTCTGTTTTCAGCAACCATCAGACCCAGGCGCCAGGCT-CTCCATCTGGCCAGTCCAC	3721
DB	966	CAGTCTCTGTTTTCAGCAACCATCAGACCCAGGCGCCAGGCTNCTCCATCTGGCCAGTCCAC	1025
QY	3722	AGCAGCAGTTCAGCCATTTTACACCGGGGCTTGGCGCAACTCCACCTCCATGACACCTG	3781
DB	1026	AGCAGCAGTTCAGCCATTTTACACCGGGGCTTGGCGCAACTCCACCTCCATGACACCTG	1085
QY	3782	CCTCCAAACACGACGTCGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGCTTTACGA	3841
DB	1086	CCTCCAAACACGACGTCGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGCTTTACGA	1145
QY	3842	TCCATCTCTTTCAGCGTTCAGCGCGGTATACCAACCCACATGCGCCACGCTACCTC	3901
DB	1146	TCCATCTCTTTCAGCGTTCAGCGCGGTATACCAACCCACATGCGCCACGCTACCTC	1205
QY	3902	AGGCTCATGTACAGTTCAGGAATGTTCTCTCATCCAACTGCCCATGCGCCCAATGATGC	3961
DB	1206	AGGCTCATGTACAGTTCAGGAATGTTCTCTCATCCAACTGCCCATGCGCCCAATGATGC	1265
QY	3962	TAATGACGACACAGCCACCCCGCGGTCCCGAGCGCGCTCCGCTCAAAGTGCACCTACAGC	4021
DB	1266	TAATGACGACACAGCCACCCCGCGGTCCCGAGCGCGCTCCGCTCAAAGTGCACCTACAGC	1325
QY	4022	CCATTCAGTCTCGACAAACAGCGCATTTCCCTATATAGCGACCCCTTCAGTACAAGCCC	4081
DB	1326	CCATTCAGTCTCGACAAACAGCGCATTTCCCTATATAGCGACCCCTTCAGTACAAGCCC	1385
QY	4082	ACCACCAACAGCAGTTGTNAGGCTGCCCTGGAGGAAACCGAAGGCGCAAAATCCCTCTCC	4141
DB	1386	ACCACCAACAGCAGTTGTNAGGCTGCCCTGGAGGAAACCGAAGGCGCAAAATCCCTCTCC	1445
QY	4142	CTTCTACTGCTTCTTCAAACTGGAAAGCAGAAAACTAGAAATTTTCAATTTTGTGTTTT	4201
DB	1446	CTTCTACTGCTTCTTCAAACTGGAAAGCAGAAAACTAGAAATTTTCAATTTTGTGTTTT	1505
QY	4202	AAAATATATATGTTGATTTCTTTGTAAATCCAAATAGGAAATGCTAACAGTTTCACTTGCAGT	4261
DB	1506	AAAATATATATGTTGATTTCTTTGTAAATCCAAATAGGAAATGCTAACAGTTTCACTTGCAGT	1565
QY	4262	GGAAGATATCTTGGACCCAGTAGAGGCATTTAGGAACTTGGGGGCTATTCCATAATTCAT	4321
DB	1566	GGAAGATATCTTGGACCCAGTAGAGGCATTTAGGAACTTGGGGGCTATTCCATAATTCAT	1625


```

;
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 232.00010101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/305-1220
; TELEFAX: 612/305-1228
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 516 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-750-323-1

Query Match      11.1%; Score 497.4; DB 8; Length 516;
Best Local Similarity 99.8%; Pred. No. 9.4e-115;
Matches 498; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

392 TTGGTAGCAACGGCAACGGCGCGCGCGTTTCGGCCCGGCTCCCGGGGCTCTCTTGGTC 451
Db 1 TTGGTAGCAACGGCAACGGCGCGCGCGTTTCGGCCCGGCTCCCGGGGCTCTCTTGGTC 60

452 TCGGGGGGCTCCCGGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 511
Db 61 TCGGGGGGCTCCCGGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120

512 CTCGGGCGCGCAACCGCGCTCCCGCTCGGGCGCGCTCGGCGCGCTCGGCGCGCTCGG 571
Db 121 CTCGGGCGCGCAACCGCGCTCCCGCTCGGGCGCGCTCGGCGCGCTCGGCGCGCTCGG 180

572 GCGTCTCTTGGCGCGCGCGCTCCCGCTCGGCGCGCTCGGCGCGCTCGGCGCGCTCGG 631
Db 181 GCGTCTCTTGGCGCGCGCGCTCCCGCTCGGCGCGCTCGGCGCGCTCGGCGCGCTCGG 240

632 GCGCCCTCAACATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 691
Db 241 GCGCCCTCAACATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 300

692 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 751
Db 301 AGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 360

752 AGCCCGCGCGCAGCGCGCTTCTAGCGTCCCGCGCGCGCGCTTCTAGCGTCCCGCGCG 811
Db 361 AGCCCGCGCGCAGCGCGCTTCTAGCGTCCCGCGCGCGCGCTTCTAGCGTCCCGCGCG 420

812 CGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 871
Db 421 CGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 480

872 GGAGGCGCGCGCTGGGCAG 890
Db 481 GGAGGCGCGCGCTGGGCAG 499

RESULT 11
US-10-956-157-3976/c
; Sequence 3976, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3976
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-956-157-3976

Query Match      10.1%; Score 452; DB 9; Length 682;
Best Local Similarity 90.5%; Pred. No. 3.3e-103;
Matches 516; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

3335 AGTCTCAGCATCTCTATAGTCTCTGTAAATACAGGGTAATCTAGTAATGATGGCAC 3394
Db 681 AGTCTCAGCATCTCTATAGTCTCTGTAAATACAGGGTAATCTAGTAATGATGGCAC 622

3395 CACCAACACAGCCCGAGCTGTTTGTAGTATCTTCTTACAGCAACTCAGTACGGGCTCATG 3454
Db 621 CACCAACACAGCCCGAGCTGTTTGTAGTATCTTCTTACAGCAACTCAGTACGGGCTCATG 562

3455 AGCAGACGATGCGATGATGATGCTCCCAAATTAACATACAAAGGAGACAAGCCCTT 3514
Db 561 AGCAGACGATGCGATGATG-----TTTCCACGGCTCCCTTTGCTCAGCAGTATGGCACCCTAACGCTA 541

3515 CTTTCTATTTCGCCATTTTCACAGGGCTCCTTGTCTCAGCAGTATGGGACCCCTAACGCTA 3574
Db 540 -----TTTCCACGGCTCCTTTGCTCAGCAGTATGGCACCCTAACGCTA 496

3575 CCCTGCACCCCATACTCCACACCCCTCAGCCTTACAGTACCCCTCAGCAGCAGCAAA 3634
Db 495 CCCTGCACCCCATACTCCACACCCCTCAGCCTTACAGTACCCCTCAGCAGCAGCAAA 436

3635 GCAACATGTTGGAGTATCTCTGCACCCAGTCTCTTTCAGCAGCAGTACAGCAGCAGCG 3694
Db 435 GCAACATGTTGGAGTATCTCTGCACCCAGTCTCTTTCAGCAGCAGTACAGCAGCAGCG 376

3695 CCCAGGCTCTCATCTGGCCAGTCCACAGCAGTACAGCAGTACAGCAGTACAGCAGCGGCTTG 3754
Db 375 CCCAGGCTCTCATCTGGCCAGTCCACAGCAGTACAGCAGTACAGCAGTACAGCAGCGGCTTG 316

3755 CGCCAACTCCACCCCTCCATGACACCTCTCCAAACAGCAGTCCGACAGCAAGATAGTTTC 3814
Db 315 CGCCAACTCCACCCCTCCATGACACCTCTCCAAACAGCAGTCCGACAGCAAGATAGTTTC 256

3815 CAGCAGCAACACAGACTGTCTTTAGCATCCATCTCTTTCAGCAGTCCAGCGGCGGTATACCA 3874
Db 255 CAGCAGCAACACAGACTGTCTTTAGCATCCATCTCTTTCAGCAGTCCAGCGGCGGTATACCA 196

3875 ACCACCCACATGGCCACGTAACCTCAGG 3904
Db 195 ACCACCCACATGGCCACGTAACCTCAGG 166

RESULT 12
US-10-956-157-9211
; Sequence 9211, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9211
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-956-157-9211

Query Match      10.1%; Score 452; DB 9; Length 682;
Best Local Similarity 90.5%; Pred. No. 3.3e-103;
Matches 516; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

3335 AGTCTCAGCATCTCTATAGTCTCTGTAAATACAGGGTAATCTAGTAATGATGGCAC 3394
Db 681 AGTCTCAGCATCTCTATAGTCTCTGTAAATACAGGGTAATCTAGTAATGATGGCAC 622

3395 CACCAACACAGCCCGAGCTGTTTGTAGTATCTTCTTACAGCAACTCAGTACGGGCTCATG 3454
Db 621 CACCAACACAGCCCGAGCTGTTTGTAGTATCTTCTTACAGCAACTCAGTACGGGCTCATG 562

3455 AGCAGACGATGCGATGATGATGCTCCCAAATTAACATACAAAGGAGACAAGCCCTT 3514
Db 561 AGCAGACGATGCGATGATG-----TTTCCACGGCTCCCTTTGCTCAGCAGTATGGCACCCTAACGCTA 541

3515 CTTTCTATTTCGCCATTTTCACAGGGCTCCTTGTCTCAGCAGTATGGGACCCCTAACGCTA 3574
Db 540 -----TTTCCACGGCTCCTTTGCTCAGCAGTATGGCACCCTAACGCTA 496

3575 CCCTGCACCCCATACTCCACACCCCTCAGCCTTACAGTACCCCTCAGCAGCAGCAAA 3634
Db 495 CCCTGCACCCCATACTCCACACCCCTCAGCCTTACAGTACCCCTCAGCAGCAGCAAA 436

3635 GCAACATGTTGGAGTATCTCTGCACCCAGTCTCTTTCAGCAGCAGTACAGCAGCAGCG 3694
Db 435 GCAACATGTTGGAGTATCTCTGCACCCAGTCTCTTTCAGCAGCAGTACAGCAGCAGCG 376

3695 CCCAGGCTCTCATCTGGCCAGTCCACAGCAGTACAGCAGTACAGCAGTACAGCAGCGGCTTG 3754
Db 375 CCCAGGCTCTCATCTGGCCAGTCCACAGCAGTACAGCAGTACAGCAGTACAGCAGCGGCTTG 316

3755 CGCCAACTCCACCCCTCCATGACACCTCTCCAAACAGCAGTCCGACAGCAAGATAGTTTC 3814
Db 315 CGCCAACTCCACCCCTCCATGACACCTCTCCAAACAGCAGTCCGACAGCAAGATAGTTTC 256

3815 CAGCAGCAACACAGACTGTCTTTAGCATCCATCTCTTTCAGCAGTCCAGCGGCGGTATACCA 3874
Db 255 CAGCAGCAACACAGACTGTCTTTAGCATCCATCTCTTTCAGCAGTCCAGCGGCGGTATACCA 196

3875 ACCACCCACATGGCCACGTAACCTCAGG 3904
Db 195 ACCACCCACATGGCCACGTAACCTCAGG 166
```

```

;
; ORGANISM: Homo sapiens
; US-10-956-157-3976

Query Match      10.1%; Score 452; DB 9; Length 682;
Best Local Similarity 90.5%; Pred. No. 3.3e-103;
Matches 516; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

3335 AGTCTCAGCATCTCTATAGTCTCTGTAAATACAGGGTAATCTAGTAATGATGGCAC 3394
Db 681 AGTCTCAGCATCTCTATAGTCTCTGTAAATACAGGGTAATCTAGTAATGATGGCAC 622

3395 CACCAACACAGCCCGAGCTGTTTGTAGTATCTTCTTACAGCAACTCAGTACGGGCTCATG 3454
Db 621 CACCAACACAGCCCGAGCTGTTTGTAGTATCTTCTTACAGCAACTCAGTACGGGCTCATG 562

3455 AGCAGACGATGCGATGATGATGCTCCCAAATTAACATACAAAGGAGACAAGCCCTT 3514
Db 561 AGCAGACGATGCGATGATG-----TTTCCACGGCTCCCTTTGCTCAGCAGTATGGCACCCTAACGCTA 541

3515 CTTTCTATTTCGCCATTTTCACAGGGCTCCTTGTCTCAGCAGTATGGGACCCCTAACGCTA 3574
Db 540 -----TTTCCACGGCTCCTTTGCTCAGCAGTATGGCACCCTAACGCTA 496

3575 CCCTGCACCCCATACTCCACACCCCTCAGCCTTACAGTACCCCTCAGCAGCAGCAAA 3634
Db 495 CCCTGCACCCCATACTCCACACCCCTCAGCCTTACAGTACCCCTCAGCAGCAGCAAA 436

3635 GCAACATGTTGGAGTATCTCTGCACCCAGTCTCTTTCAGCAGCAGTACAGCAGCAGCG 3694
Db 435 GCAACATGTTGGAGTATCTCTGCACCCAGTCTCTTTCAGCAGCAGTACAGCAGCAGCG 376

3695 CCCAGGCTCTCATCTGGCCAGTCCACAGCAGTACAGCAGTACAGCAGTACAGCAGCGGCTTG 3754
Db 375 CCCAGGCTCTCATCTGGCCAGTCCACAGCAGTACAGCAGTACAGCAGTACAGCAGCGGCTTG 316

3755 CGCCAACTCCACCCCTCCATGACACCTCTCCAAACAGCAGTCCGACAGCAAGATAGTTTC 3814
Db 315 CGCCAACTCCACCCCTCCATGACACCTCTCCAAACAGCAGTCCGACAGCAAGATAGTTTC 256

3815 CAGCAGCAACACAGACTGTCTTTAGCATCCATCTCTTTCAGCAGTCCAGCGGCGGTATACCA 3874
Db 255 CAGCAGCAACACAGACTGTCTTTAGCATCCATCTCTTTCAGCAGTCCAGCGGCGGTATACCA 196

3875 ACCACCCACATGGCCACGTAACCTCAGG 3904
Db 195 ACCACCCACATGGCCACGTAACCTCAGG 166

RESULT 12
US-10-956-157-9211
; Sequence 9211, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9211
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-956-157-9211

Query Match      10.1%; Score 452; DB 9; Length 682;
Best Local Similarity 90.5%; Pred. No. 3.3e-103;
Matches 516; Conservative 0; Mismatches 0; Indels 54; Gaps 1;

3335 AGTCTCAGCATCTCTATAGTCTCTGTAAATACAGGGTAATCTAGTAATGATGGCAC 3394
Db 681 AGTCTCAGCATCTCTATAGTCTCTGTAAATACAGGGTAATCTAGTAATGATGGCAC 622

3395 CACCAACACAGCCCGAGCTGTTTGTAGTATCTTCTTACAGCAACTCAGTACGGGCTCATG 3454
Db 621 CACCAACACAGCCCGAGCTGTTTGTAGTATCTTCTTACAGCAACTCAGTACGGGCTCATG 562

3455 AGCAGACGATGCGATGATGATGCTCCCAAATTAACATACAAAGGAGACAAGCCCTT 3514
Db 561 AGCAGACGATGCGATGATG-----TTTCCACGGCTCCCTTTGCTCAGCAGTATGGCACCCTAACGCTA 541

3515 CTTTCTATTTCGCCATTTTCACAGGGCTCCTTGTCTCAGCAGTATGGGACCCCTAACGCTA 3574
Db 540 -----TTTCCACGGCTCCTTTGCTCAGCAGTATGGCACCCTAACGCTA 496

3575 CCCTGCACCCCATACTCCACACCCCTCAGCCTTACAGTACCCCTCAGCAGCAGCAAA 3634
Db 495 CCCTGCACCCCATACTCCACACCCCTCAGCCTTACAGTACCCCTCAGCAGCAGCAAA 436

3635 GCAACATGTTGGAGTATCTCTGCACCCAGTCTCTTTCAGCAGCAGTACAGCAGCAGCG 3694
Db 435 GCAACATGTTGGAGTATCTCTGCACCCAGTCTCTTTCAGCAGCAGTACAGCAGCAGCG 376

3695 CCCAGGCTCTCATCTGGCCAGTCCACAGCAGTACAGCAGTACAGCAGTACAGCAGCGGCTTG 3754
Db 375 CCCAGGCTCTCATCTGGCCAGTCCACAGCAGTACAGCAGTACAGCAGTACAGCAGCGGCTTG 316

3755 CGCCAACTCCACCCCTCCATGACACCTCTCCAAACAGCAGTCCGACAGCAAGATAGTTTC 3814
Db 315 CGCCAACTCCACCCCTCCATGACACCTCTCCAAACAGCAGTCCGACAGCAAGATAGTTTC 256

3815 CAGCAGCAACACAGACTGTCTTTAGCATCCATCTCTTTCAGCAGTCCAGCGGCGGTATACCA 3874
Db 255 CAGCAGCAACACAGACTGTCTTTAGCATCCATCTCTTTCAGCAGTCCAGCGGCGGTATACCA 196

3875 ACCACCCACATGGCCACGTAACCTCAGG 3904
Db 195 ACCACCCACATGGCCACGTAACCTCAGG 166
```

Db 2 AGTCTCAGCATCCTCATGTCTATAGTCTGTAATAACAGGGTAAGTGAATGATGGCAC 61
Qy 3395 CACCAACACAGCCAGCCTGTTAGTATCTTCTCAGCAACTCAGTACGGGGTCTATG 3454
Db 62 CACCAACACAGCCAGCCTGTTAGTATCTTCTCAGCAACTCAGTACGGGGTCTATG 121
Qy 3455 AGCAGACCATGCGATGATGTCCTCCAAATTACCATACAACAGGAGAGAAAGCCCTT 3514
Db 122 AGCAGACCATGCGATGATG-----142
Qy 3515 CTTTCTACTTTTCCATTTTCCAGGGCTCCCTTGCTCAGCAGTATGCGACCTTAACGCTA 3574
Db 143 -----TTTCCAGGGCTCCCTTGCTCAGCAGTATGCGACCTTAACGCTA 187
Qy 3575 CCCTGCACCCACATCTCCACACCTCAGCCTTTCAGCTTACCCCACTGGACAGCAGCAAA 3634
Db 188 CCCTGCACCCACATCTCCACACCTCAGCCTTTCAGCTTACCCCACTGGACAGCAGCAAA 247
Qy 3635 GCCAATATGTTGGAAGTATCTCTGACCCAGTCTCTGTTTCAGACCATCAGCACAGGCGG 3694
Db 248 GCCAATATGTTGGAAGTATCTCTGACCCAGTCTCTGTTTCAGACCATCAGCACAGGCGG 307
Qy 3695 CCCAGGCTCTCCATCTGSCCAGTCCACAGCAGCAGTACGCCATTTACCAAGGGGCTTG 3754
Db 308 CCCAGGCTCTCCATCTGSCCAGTCCACAGCAGCAGTACGCCATTTACCAAGGGGCTTG 367
Qy 3755 GCCAATCTCCACCTTCCATGACCTGCTCCACACAGCAGTCCGACAGAAATAGTTTCC 3814
Db 368 GCCAATCTCCACCTTCCATGACCTGCTCCACACAGCAGTCCGACAGAAATAGTTTCC 427
Qy 3815 CAGCAGCAACAAGACTCTTTTACGATCCATCTTTCAGCTTACGGCGCGTATACCA 3874
Db 428 CAGCAGCAACAAGACTCTTTTACGATCCATCTTCTCTCAGTTTACGGCGCGTATACCA 487
Qy 3875 ACCACCCCATGCGCCCATCTACTCAGG 3904
Db 488 ACCACCCCATGCGCCCATCTACTCAGG 517
RESULT 13
US-10-363-345A-23185/c
; Sequence 23185, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 23185
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-Island No: 23185
US-10-363-345A-23185
Query Match 9.8%; Score 441.2; DB 8; Length 704;
Best Local Similarity 85.0%; Pred. No. 1.8e-100;
Matches 505; Conservative 0; Mismatches 88; Indels 1; Gaps 1;
Qy 1 ACCCCCGAGAAAGCAACCCAGCGCGCCGCTCCTCAGCTGTCCCTCCCGGCCCGGG 60
Db 112 ACCCCCGAGAAAGCAACCAACGCGCGCCGCTCCTCAGCTATCCCTCCCGACCCCGAA 171
Qy 61 GCCACCTCAGCTTCTGCTTCGCTCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120
Db 593 ACCCCCGAGAAAGCAACCAACGCGCGCCGCTCCTCAGCTATCCCTCCCGACCCCGAA 534
Qy 61 GCCACCTCAGCTTCTGCTTCGCTCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120

Db 533 ACCACCTCAGCTTCTACTTCCGTCTAACCCCTCCGACTTCCGATAAAAAATCCCTATCCG 474
Qy 121 CACCTTCGGCTCCCAACCCGGCGCTTCGGCGCCCGCCCTCCGATGCGCTCAGCGGCCGCA 180
Db 473 CACCTTCGGCTCCCAACCCGGCGCTTCGAGCGCGCCCGCTCCGATGAGCTCAACGACCGCA 414
Qy 181 GCTCTTCGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTCGCGCAGCAGAGTGGGCC 240
Db 413 ACTCTCGAAATCCCGGATAAACCAACGAAATCTGCGG-TTCGCGCAACCAAAATAACCC 355
Qy 241 GGGTGGCGCTCGCTCCAGCGCGCGCGGCGGAGCGGGGGGGGGGGGGGGGGGGGG 300
Db 354 GAATAACGCTCGCTCCAAACGACCGAGCGAGAAACGAAACGAGATAACGCGGACC 295
Qy 301 CCGGAGACGATATCCCTCCCGCGCCCTCCCGCGCGCGCGCCCGCCCTCCCTCCCTCCCGG 360
Db 294 CCGAAACGATATCCCTCCCGCGCCCTCCCTCCCGCGCGCGCGCCCGCTCCCTCCCGA 235
Qy 361 CAGAGTCGCTCCCTCCCGCTCAGACTGTTTGGTAGCAACGGAACGCGCGCGCGCG 420
Db 234 CAAAACTCGCTCCCTCCCGCTCAAACTATTTTAAATAACAAACGACAAACGAGACGCG 175
Qy 421 TTTCCGCGCGCGCTCCCGCGCGCTCTTGGTCTCGCGCGCGCGCTCCCGCGCTTCGTGTC 480
Db 174 TTTCCGACCGGACTCCCGGAGCTCTTAAATCTCGACGAACCTCCCGCGCGCTTCGTGTC 115
Qy 481 GTCTCTTCTCCCGCTCGCAGCGCGCGCGCGCTCCCGCGCGCGCGCGCGCTCCCGG 540
Db 114 CTCTCTTCTCCCGCTCGCAACCGCGAAACGCGCTCCGACCGCGCGCGCGCGCTCCCG 55
Qy 541 CTCGGCGCGCTGCTCCCGCGCGCTTCGGGCTCTCTTTGGCGCGCGCGCT 594
Db 54 CTCGACGCGCGCGTCCCGCGCGTTCGACGCTCTTAAACGCGCGCGGACT 1
RESULT 14
US-10-363-345A-23186
; Sequence 23186, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 23186
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-Island No: 23186
US-10-363-345A-23186
Query Match 9.8%; Score 441.2; DB 8; Length 704;
Best Local Similarity 85.0%; Pred. No. 1.8e-100;
Matches 505; Conservative 0; Mismatches 88; Indels 1; Gaps 1;
Qy 1 ACCCCCGAGAAAGCAACCCAGCGCGCCGCTCCTCAGCTGTCCCTCCCGGCCCGGG 60
Db 112 ACCCCCGAGAAAGCAACCAACGCGCGCCGCTCCTCAGCTATCCCTCCCGACCCCGAA 171
Qy 61 GCCACCTCAGCTTCTGCTTCGCTCTGACCCCTCCGACTTCGGTAAAGAGTCCCTATCCG 120
Db 172 ACCACCTCAGCTTCTACTTTCGCTCTAACCCCTCCGACTTCGATATAAAAAATCCCTATCCG 231
Qy 121 CACCTTCGGCTCCCAACCCGGCGCTTCGGCGCGCGCGCTCCGATGCGCTCAGCGCGCGCA 180
Db 232 CACCTTCGGCTCCCAACCCGGCGCTTCGACGCGCGCGCTCCGATACGCTCAACGACCGCA 291

```
QY 181 GTCCTCTGAGTCCCGCGTGCCACCGAGTCTGCGCGTTCGCGCAGCCAGGTGGCCC 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
292 ACTCTCGAAATCCCGGATTAACCAACCGAATCTGGCG-TTCGCGCAACCAATAACCC 350
QY 241 GSGTGGCGTCCGTCAGCGCGCGCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCG 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
351 GAATAACGCTCGTCCCAACGACGACGACGACGACGACGACGACGACGACGACGACG 410
QY 301 CGGGACCGTATCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
411 CGGAACCGTATCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 470
QY 361 CAGAGCTCGCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
471 CAAAACCTCGCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 530
QY 421 TTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
531 TTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 590
QY 481 GTCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
591 CTCCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 650
QY 541 CTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
651 CTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 704
```

RESULT 15

```
US-10-363-483A-23185/c
; Sequence 23185, Application US/10363483A
; Publication No. US2005006401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 23185
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 23185
US-10-363-483A-23185
```

```
Query Match 9.8%; Score 441.2; DB 9; Length 704;
Best Local Similarity 85.0%; Pred. No. 1.8e-100;
Matches 505; Conservative 0; Mismatches 88; Indels 1; Gaps 1;
```

```
QY 1 ACCCCGAGAAAGCAACCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 60
DB 593 ACCCCGAGAAAGCAACCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 534
QY 61 GCCACCTCAGCTTCTGCTTCCGCTGACCCCTCCGACTTCCGGTAAAGAGTCCCTATCCG 120
DB 533 ACCACCTCAGCTTCTGCTTCCGCTTACCCCTCCGACTTCCGATATAAATCCCTATCCG 474
QY 121 CACCTCCGCTCCCAACCGCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCA 180
DB 473 CACCTCCGCTCCCAACCGCGCGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCA 414
QY 181 GCTCCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGCGCGCGCGCGCGCGCG 240
DB 413 ACTCCTCGAAATCCCGCGATACCAACCGAATCTCGCGG-TTCGCGCGCAACCAATAACCC 355
```

```
QY 241 GSGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
354 GAATAACGCTCGCTCCAAACGACGCGACGCGACGAAACGAAACGAAACGAGATAACGCGACC 295
QY 301 CGGGACCGTATCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
294 CGGAACCGTATCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 235
QY 361 CAGAGCTCGCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
234 CAAAACCTCGCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 175
QY 421 TTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
174 TTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 115
QY 481 GTCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
114 CTCCTTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 55
QY 541 CTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
54 CTCGAGCGCGCGCGCGTCCCGCGCGCGTTCGAGCGTCTCCTTAACGCGCGCGCGACT 1
```

Search completed: December 23, 2005, 06:06:39
Job time : 3078 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2005, 18:44:03 ; Search time 944 Seconds
(without alignments)
2463.386 Million cell updates/sec

Title: US-10-802-228-1
Perfect score: 4481
Sequence: 1 acccccgagaaagcaaccaccc.....taaaaaaaaaaaaaaaaaaaaaa 4481

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4169288 seqs, 259477437 residues

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New:
1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81.4	1.8	13672	7	US-11-055-035-2
2	80.6	1.8	255	6	US-10-909-125-1963
c 3	70.2	1.6	5679	7	US-11-075-185-36
c 4	70.2	1.6	78869	7	US-11-075-185-1
5	69.6	1.6	1867	7	US-11-043-590-45
6	69.2	1.5	1310	7	US-11-137-671-13
7	69.2	1.5	3263	7	US-11-137-671-15
8	69.2	1.5	4286	7	US-11-137-671-14
c 9	69.2	1.5	11070	7	US-11-075-185-34
10	68.4	1.5	3507	6	US-10-509-422-1
11	68.4	1.5	3704	6	US-10-509-422-3
12	68.4	1.5	20945	6	US-10-995-561-13463
13	68.4	1.5	23082	6	US-10-995-561-13457
c 14	68.4	1.5	94510	6	US-10-995-561-13332
c 15	67.4	1.5	14248	6	US-10-995-561-13381
c 16	67.4	1.5	96128	6	US-10-995-561-13197
17	66.8	1.5	2276	6	US-10-131-826A-9
c 18	66.4	1.5	16082	6	US-10-995-561-13485
c 19	66.4	1.5	23894	6	US-10-995-561-13320
c 20	66.4	1.5	46215	6	US-10-995-561-13483
c 21	66.4	1.5	268685	6	US-10-933-025-22
c 22	65.8	1.5	171936	6	US-10-933-025-24
c 23	65.6	1.5	1433	7	US-11-090-351-1

c 24	65	1.5	12777	6	US-10-477-507A-3	Sequence 3, Appli
c 25	64.8	1.4	1357	6	US-10-996-217A-6	Sequence 6, Appli
c 26	64.8	1.4	172147	7	US-11-112-908-22	Sequence 22, Appli
c 27	64.8	1.4	188682	7	US-11-112-908-23	Sequence 23, Appli
c 28	64.4	1.4	191684	7	US-11-121-086-2	Sequence 2, Appli
c 29	64.2	1.4	126552	7	US-11-121-086-1	Sequence 1, Appli
c 30	64	1.4	180654	7	US-11-121-086-58	Sequence 58, Appli
c 31	63.4	1.4	833	6	US-10-750-185-56670	Sequence 56670, A
c 32	63.2	1.4	167891	7	US-11-121-086-14	Sequence 14, Appli
c 33	63	1.4	170995	7	US-11-121-086-35	Sequence 35, Appli
c 34	62.6	1.4	3108	6	US-10-477-507A-1	Sequence 1, Appli
c 35	62.6	1.4	23983	6	US-10-995-561-13491	Sequence 13491, A
c 36	62.4	1.4	592	6	US-10-632-150-31	Sequence 31, Appli
c 37	62.4	1.4	592	7	US-11-073-457-31	Sequence 31, Appli
c 38	62.4	1.4	592	7	US-11-073-460-31	Sequence 31, Appli
c 39	61.8	1.4	1546	8	US-11-112-944-14	Sequence 14, Appli
c 40	61.6	1.4	165857	7	US-11-121-086-34	Sequence 34, Appli
c 41	61.4	1.4	1412	7	US-11-186-284-154	Sequence 154, App
c 42	61.4	1.4	4860	6	US-10-971-982-1	Sequence 1, Appli
c 43	61.4	1.4	127340	7	US-11-112-908-35	Sequence 35, Appli
c 44	61.2	1.4	1425	7	US-11-061-869-6	Sequence 6, Appli
c 45	61.2	1.4	2023	6	US-10-995-561-55	Sequence 55, Appli

ALIGNMENTS

RESULT: 1
US-11-055-035-2
; Sequence 2, Application US/11055035
; Publication No. US20050256072A1
; GENERAL INFORMATION:
; APPLICANT: ARONIN, NEIL
; APPLICANT: ZAMORE, PHILLIP D.
; APPLICANT: BRODERICK, JENNIFER
; TITLE OF INVENTION: DUAL FUNCTIONAL OLIGONUCLEOTIDES FOR USE IN REPRESSING
; TITLE OF INVENTION: MUTANT GENE EXPRESSION
; FILE REFERENCE: UMY-095
; CURRENT APPLICATION NUMBER: US/11/055,035
; CURRENT FILING DATE: 2005-02-09
; PRIOR APPLICATION NUMBER: 60/543,467
; PRIOR FILING DATE: 2004-02-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 2
; LENGTH: 13672
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-055-035-2

Query Match	1.8%	Score 81.4; DB 7; Length 13672;
Best Local Similarity	57.5%;	Pred. No. 1.6e-07;
Mismatches	165;	Conservative 0; Mismatches 121; Indels 1; Gaps 1;
Qy	557	CCCGCGCGGTTCGGCGTCTCTTGGCGGCGCGGCTCCCGGTGTCCCGCGCGCGGT 616
Db	263	CCGCGAGTCGGCGCGGAGCTCCGCGGACTCCGCGGCGGAGACCGCATGCGCA 322
Qy	617	GCAGCGGTGTATGGGCGCTTCCATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGC 676
Db	323	CCCTGGAAGCTGATGAAGCGCTTCAAGTCTTCCAGCAGCAGCAGCAGCAGCAGC 382
Qy	677	AGCAGCAACAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 736
Db	383	AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 441
Qy	737	CTGCCAATGTCGCAAGCCCGCGGCGGCGGCTTCTAGGTCGCGCGCGCGCGCTT 796
Db	442	CCGCGCGCGCGCGCGCGCGCTTCTCAGCTTCTCAGCGCGCGCGCGCGCAGCAGC 501
Qy	797	CGCGGTCTCGTCTCGGTCTCTCTCGTCCTCGGCGCAGCGGTCTCCCTCC 843
Db	502	CTGCTGCTCAGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCACC 548

Query Match	Best Local Similarity	Score	DB 7	Length	DB 8	Indels	Gaps	0
3208	473	3149	533	3089	593	3029	653	2969
3089	593	3029	653	2969	713	2909	772	2849
832	2789							
RESULT 10								
US-10-509-422-1								
; Sequence 1, Application US/10509422								
; Publication No. US2005024825A1								
; GENERAL INFORMATION:								
; APPLICANT: Liou, Simon								
; TITLE OF INVENTION: Human BMP2 Inducible Kinases								
; FILE REFERENCE: 004974.01015								
; CURRENT APPLICATION NUMBER: US/10/509,422								
; CURRENT FILING DATE: 2004-09-24								
; PRIOR APPLICATION NUMBER: PCT/EP03/080825								
; PRIOR FILING DATE: 2003-03-20								
; PRIOR APPLICATION NUMBER: US 60/367,512								
; PRIOR FILING DATE: 2002-03-27								
; PRIOR APPLICATION NUMBER: US 60/406,936								
; PRIOR FILING DATE: 2002-08-30								
; NUMBER OF SEQ ID NOS: 9								
; SOFTWARE: FastSeq for Windows Version 4.0								
; SEQ ID NO 1								
; LENGTH: 3507								
; TYPE: DNA								
; ORGANISM: Homo sapiens								
US-10-509-422-1								
Query Match 1.5%; Score 68.4; DB 6; Length 3507;								
Best Local Similarity 92.3%; Pred. No. 5.1e-05;								
Matches 72; Conservative 0; Mismatches 6; Indels 0; Gaps 0								
653	1379	713	1439					
RESULT 11								
US-10-509-422-3								
; Sequence 3, Application US/10509422								
; Publication No. US20050244825A1								

[illegible]

;; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
;; TITLE OF INVENTION: DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001559
;; CURRENT APPLICATION NUMBER: US/10/995,561
;; CURRENT FILING DATE: 2004-11-24
;; NUMBER OF SEQ ID NOS: 85702
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13332
;; LENGTH: 94510
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)...(94510)
;; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-995-561-13332

Query Match 1.5%; Score 68.4; DB 6; Length 94510;
Best Local Similarity 49.4%; Pred. No. 0.00023;
Matches 175; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 538 CCCTCCGCGCCGCTCCCGCGCGCTTCCTTGGCGGCCCGGCTCCC 597
DB 6351 CCCTCCCGCTCCGCGCTCCCGCGCGCTTGTGGTAGGGGAGCAGAGCCG 6292

QY 598 GGCTGTCCCGCGCGCTCGAGCGGTATATGGCCCTCACCATGTGCTGAAGCCC 657
DB 6291 AGTCAGAGACCTTTTGACGACCTTGGCGCGGAGCAGCGGATCAGCCGCTGCCGCGC 6232

QY 658 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 717
DB 6231 AAGATGCTGGAGCYGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 6172

QY 718 CAGCAGCGCGCGCGCGCGCTGCCAATGTC---CGCAAGCCCGCGCGCGCGCTTCTA 774
DB 6171 GAGAGGCGCGCTCGGGAGGCCATGGCGGCGCGGCGCTCCGCGCGCGCGCGCGC 6112

QY 775 GCCTGCGCCCGCGCGCGCTTCGCGCTTCCTGCTCGCTTCCTGCTCGCTCGCGCACG 834
DB 6111 GCTCCCG 6052

QY 835 GCTCCCTCTCGGTGTCGCGCGACCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 888
DB 6051 GCGCGGCGCTCAGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5998

RESULT 15
US-10-995-561-13381/c
; Sequence 13381, Application US/10995561
; Publication No. US2005072054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13381
; LENGTH: 14248
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13381

Query Match 1.5%; Score 67.4; DB 6; Length 14248;
Best Local Similarity 52.2%; Pred. No. 0.00016;
Matches 180; Conservative 0; Mismatches 156; Indels 9; Gaps 1;

QY 23 GCGCGCGCGCTCCTCAGGTGTCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 82
DB 5892 GCGCGCGCGCGCGCGCGCGCTCCCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5833

Search completed: December 23, 2005, 06:22:29
Job time : 948 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame plus_p2n model
Run on: December 23, 2005, 05:02:12 ; Search time 13935 Seconds
(without alignments)
5351.892 Million cell updates/sec

Title: US-10-802-228-2
Perfect score: 6961
Sequence: 1 MRSAAAAPSPAVATESRRF.....TAHFPPYMTSPVQAHHQOOL 1312

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n_model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10802228/runat_21122005_110224_9699/app_query.fasta_1.1479
-DB=genembl -Qfmt=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOCL=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10802228@cgn 1.10508 @runat_21122005_110224_9699 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb_ba.*
2: gb_in.*
3: gb_env.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pr.*
9: gb_ro.*
10: gb_ets.*
11: gb_ey.*
12: gb_un.*
13: gb_vi.*
14: gb_htg.*
15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6961	100.0	4481	6	AR153580 Sequence
2	6961	100.0	4481	6	AR401379 Sequence
3	6961	100.0	4481	6	AR447280 Sequence

4	6961	100.0	4481	6	AR632098	AR632098 Sequence
5	6961	100.0	4481	6	AX392465	AX392465 Sequence
6	6961	100.0	4481	8	HSU70323	U70323 Human ataxi
7	6890	99.0	4479	6	CQ727533	CQ727533 Sequence
8	6377	91.6	4163	8	HSDANSCA2	Y08262 H.sapiens m
9	6377	91.6	4200	6	A62706	A62706 Sequence 7
10	6025.5	86.6	4225	9	AF041472	AF041472 Mus muscu
11	5606.5	80.5	3798	6	AR632099	AR632099 Sequence
12	2612	37.5	2558	5	BC097692	BC097692 Xenopus l
13	2380	34.2	2766	8	AK095017	AK095017 Homo sapi
14	1876	27.0	1257	6	AR447281	AR447281 Sequence
15	1876	27.0	1257	6	AR632110	AR632110 Sequence
16	1876	27.0	1257	6	AX392475	AX392475 Sequence
17	1876	27.0	1257	9	MMU70870	U70870 Mus musculu
18	1876	27.0	3781	8	HS3117973	AJ3117973 Homo sapi
19	1846	26.5	3758	8	HS3117974	AJ3117974 Homo sapi
20	1846	26.5	3893	8	HS3117972	AJ3117972 Homo sapi
21	1846	26.5	3897	8	HS3117971	AJ3117971 Homo sapi
22	1846	26.5	4380	8	HS3117970	AJ3117970 Homo sapi
23	1832	26.3	4283	8	BC068012	BC068012 Homo sapi
24	1819	26.1	3635	8	BC082760	BC082760 Homo sapi
25	1819	26.1	3764	9	BC054483	BC054483 Mus muscu
26	1816.5	26.1	4674	8	AF034373	AF034373 Homo sapi
27	1780.5	25.6	3920	8	AY188335	AY188335 Homo sapi
28	1775.5	25.5	3980	8	AY188337	AY188337 Homo sapi
29	1773.5	25.5	4272	8	AY188334	AY188334 Homo sapi
30	1768.5	25.4	4012	8	AY188336	AY188336 Homo sapi
31	1768.5	25.4	4496	8	AY188338	AY188338 Homo sapi
32	1481	21.3	1080	9	MMZ78162	Z78162 M.musculus
33	1426	20.5	168247	8	AC005014	AC005014 Homo sapi
34	1310	18.8	3486	9	BC043451	BC043451 Mus muscu
35	1305	18.7	3570	5	BC057536	BC057536 Danio rer
36	1273	18.3	107717	8	AC137055	AC137055 Homo sapi
37	1257	18.1	4369	6	CQ851373	CQ851373 Sequence
38	1257	18.1	4369	8	AK128613	AK128613 Homo sapi
39	1245.5	17.9	108139	14	AC004085_3	Continuation (4 of
40	1222	17.6	2809	9	BC034083	BC034083 Mus muscu
41	1151	16.5	182703	14	AC073117	AC073117 Homo sapi
42	1012.5	14.5	623	6	AR159546	AR159546 Sequence
43	907	13.0	572	6	AR159558	AR159558 Sequence
44	857	12.3	516	6	AR447279	AR447279 Sequence
45	857	12.3	516	6	AR632097	AR632097 Sequence

ALIGNMENTS

RESULT 1
AR153580
LOCUS AR153580 AR153580 4481 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 18 from patent US 6235872.
ACCESSION AR153580
VERSION AR153580.1 GI:15121112
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4481)
AUTHORS Bredesen,D.E. and Rabizadeh,S.
TITLE Proapoptotic peptides dependence polypeptides and methods of use
JOURNAL Patent: US 6235872-A 18 22-MAY-2001;
FEATURES Location/Qualifiers
source
1. .4481
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2.26e-97 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-802-228-2 (1-1312) x AR153580 (1-4481)

Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 163 ATGCGCTCAGCGCGCGCAGCTCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTC 222
Qy 21 AlaAlaAlaArgTrrProGlyTrrArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db 223 GCCGAGCCAGGTGGCGCGGCTCGCTCCAGCGCGCGCGCGCGCGAGCGGGCGG 282
Qy 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 283 GCGCGCGTGGCGCGCGCGCGGACGATATCCCTCCGCGCGCCCTCCCGCGCGCGCC 342
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db 343 GCGCGCGCTCCCTCCCGCGCAGAGCTCGCTCCCTCGCGCTCAGACTGTTTTGGTAGCAAC 402
Qy 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgLeuLeuGlyGlyGlyPro 100
Db 403 GCGAAACGGCGCGCGCGCTTTCGGCGCGGCTCCCGCGCGCTCCCTTGGTCTCGCGCGGCT 462
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
Db 463 CCGCGCGCTTCGTGCTGCTCTTCGCCCTTCGCCAGCGCGCGCGCGCGCTCCCGCGCG 522
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 523 CCAACCGCGCTCCCGCTCGCGCGCGGTGGTCCCGCGCGGTTCGGCGCTCTCTTG 582
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 583 GCGCGCGCGCTCCCGCTGCTCCCGCGCGCGGTGCGAGCGGTGTATGGCGCGCTCAC 642
Qy 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 643 ATGTGCTGAACCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
Qy 181 GlnGlnGlnGlnGlnGlnProProAlaAlaAlaAsnValArgLysProGlyGly 200
Db 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 762
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
Db 763 AGCGGCTTCTAGCGTGGCGCGCGCGCTTCGCGCTCTCGCTCTCGCTCTCTCTCG 822
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
Db 823 TCCTCGCGCCACGGCTCCCTCCCTCGGTGGTTCGCGGACCTCCGCGCGCGCGCGCG 882
Qy 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
Db 883 CTGGCGAGGGTTCGAAACAGTAACAGGACTGCTCAGTCTACGATTTCTTTTATGGA 942
Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
Db 943 ATCTATGCAAAATATGAGGATGGTTATATATCATCATCATGTTGTTGGTCCAAATGTGA 1002
Qy 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db 1003 GTACAAGTGAAAAATGGAGGTATATGAAGGAGTTTTTAAACTTACAGTCCGAAGTGT 1062
Qy 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
Db 1063 GATTTGGTACTTTGATGCCGCATCGAATAAGTACAGATTCAGTTCGGGGCCGAAACGT 1122
Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
Db 1123 GAAGAAATAATGGAGATATTTTGTTCAAATGTTTCAGACTTTGTTGTGTACAGTTTAA 1182
Qy 341 AspMetAspSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360

Db 1183 GATATGAGCTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA 1242
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTrrPaspAlaGlyGluLeuThrAla 380
Db 1243 GTGAATGGCGGAACACAAAGAGAGAGCTGGAGCCCTGGGATGCAGGTGAACACTCACAGCC 1302
Qy 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrrPaspProAsnAspMet 400
Db 1303 AATGAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGATGGATGCCCAATGATG 1362
Qy 401 PheArgTyrAsnGluGluAsnTyrGlyValSerThrTyrAspSerSerLeuSerSer 420
Db 1363 TTTTGATATATGAAGAAAAATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db 1423 TATACAGTGCCCTTAGAAAGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGCA 1482
Qy 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db 1483 AACCAAGTTAGCAGAAGAAATTTAGTCAAGTCCCGCAGTACAAAGCTCGAGTGGCCCTG 1542
Qy 461 AsnAspAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db 1543 AATGATGATAGAGTGGAGAGAAAAATACACACAGTTTCAGAGAAATTTCCAGTGAACGT 1602
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
Db 1603 GAGGGCACAGCATAAACACTAGGGAAAAATAATATATATCTCTCTGGCAAAAGAAATAGA 1662
Qy 501 GluValIleSerTrrPrrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1663 GAAGTCATATCTCGGGAAAGTGGGAGACAGAAATTCACCGGTATGGGCGAGCTGGATCG 1722
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATTCGTGTTCA 1782
Qy 541 AspGlnArgValValAsnGlyGlyValProTrrPrrProSerProCysProSerProSerSer 560
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCTTGCCTCATCTCTCTCTCT 1842
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1843 CGCCACCTTCTCGCTACCACTCAGTCAGGTCCCAACTCTCTTCCACCTCGGCGCAGCCCT 1902
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db 1903 ACACGGCGCGCTCCAGCGCCCTCCGCGGCATCCAGACCCCTCTCACCCCTCTGCT 1962
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1963 CATGTTCTCCAGTCTCTGCTCTACTATGCTTAAACGCATGCTTTCAGAAAGGCGCTCCA 2022
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Db 2023 AGGATGTCCCAAGGCGCAGCAGCATCTCGAAATCACAGAGTTTCTGCTGGGAGGGT 2082
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
Db 2083 TCCATATCCAGTGGCTTAGAAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCT 2142
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTrrPrrSerSerValSerGlyValPro 680
Db 2143 CCAGTAGCAAGACCACTCCCTCCGGGGGAACGTGGTCACTCAGTGGTTCAGTGGGCTCCA 2202
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db 2203 AGATTATCCCTAAACACTCATAGACCCAGGTCTCCAGACAGAAACAGATTATGGAAATACC 2262
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
Db 2263 CCCAGTGGCGCGATTCTTGCTTCTCCCAAGCTGGTATTTATTTCAACTGAAGCTGTGCC 2322

```
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGCCTATTCCAGCTGCATCTCTACGCTGTAGTCTCTGCATCGAACAGAGCTGTATCC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2383 CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGAGAACTCTCTCTGAGGG 2442
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2443 AATAAGAAAATATTAACCCCAATGAACATCACCTAGTCTCTCAAAAGCTGAACAAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
DB 2503 GGTATATCACCACTGTTCTGAACATAGAAAACAGATTGATTTAAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 AATGATTTTAGGTTAGAGCCAGTTCTACTTCTGAATCTATGGATCAACTACTTAAACAAA 2622
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2623 AATAGAGAGGAGAAAATCAAGAGATTGTATCAAGACAAAATTTGAAACAAGTGTAG 2682
QY 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
DB 2683 GATTTCTTTCATTTGAAATAGCAGCAGCAACTGTATACAGTGGCAGCAGCAAGCCGATAGC 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisArgGlyProGluValThr 880
DB 2743 CCAGCAATTTCCCTTTCAATACTTAGTAACACGAGCACAAGAGGGAGCTGTAGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
DB 2803 TCCCAAGGGTTCCAGCTTCCAGCCCAAGCATGTAAACAGAGAAAACGATTAAGGAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrIleuAsnProAsnAlaLysGluPhe 920
DB 2863 AAGAAAGACGACGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAGGAGTTT 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrProArgProGln 940
DB 2923 AACCCAGTTTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCGCTCAA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
DB 2983 GCACAACTTAGCCCATCTATGGTGGGTTCATCAACAGCCAACTCCAGTTTATCTCAGCCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
DB 3043 GTTTGTTTGCACCAATATGATGATCCAGTCCAGTCCAGTCCAGCGGCGTCAACCTTTA 3102
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
DB 3103 TACCAATACCTATGACGCCCATGCCAGTCAATCAAGCCCAAGACATATAGCAGTACCA 3162
QY 1001 AsnMetProGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
DB 3163 AATATGCCCCAACAGCGGCAAGACCAAGCATCATCAGATGCCATGATGACCCAGCGTCA 3222
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
DB 3223 GCAGGGGCCCCCGATTGAGCAGCACCCACCGAGTTACTCCAGCAATATGTTGCCATC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
DB 3283 AGTCCTCAGCAGTTTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATATTCAGTCTCAG 3342
QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
DB 3343 CATCCTCATGTCTATAGTCTCTGTAAATCAGGGTAATGCTAGAATGATGGCACCACCAACA 3402
```

```
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
DB 3403 CACGCCCCAGCTGGTTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGCG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
DB 3463 CATGCGATGTATGCATGTCCTCCAAATTTACCATACAACAGGAGACAAGCCCTTCTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
DB 3523 TTTGCAATTTCCAGGGCTCCCTTGTCTCAGCAGTATGCGCACCTAACCTACCTTGCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
DB 3583 CCACATATCTCCACACCTCAGCTTCCAGTCTCCCTACCTTGGACAGCAGCAAGCCCAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
DB 3643 GGTGGAAGTCTCTGCTGCTCAGCAGTCTTACCGCGGGCTTGCAGCAACT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCACTTACCGCGGGCTTGCAGCAACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
DB 3763 CCACCTCTCCATGACACCTGCTCCCAACAGCAGTCCGACAGCAATAGTTTCCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnPro 1240
DB 3823 CAACGACATGCTTTTACGATCCATCTTCTCAGTTCAGCGCGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3883 CACATGGCCCACTACCTCAGGCTCATGTACAGTACAGTACAGCAATGTTCTTCTCATCTCA 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaLeu 1280
DB 3943 GCCCATGCCCCAATGATGCTAATGACGACACAGCCACCCCGGGGTCCCAGGCGGCCCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
DB 4003 GCTCAAGTGCATACAGCCCATTCAGTCTCGACAACAGCGCATTTCCCTATATGACG 4062
QY 1301 HisProSerValGlnAlaHisGlnGlnLeu 1312
DB 4063 CACCCTTCAGTACAGCCCAACCCACCAACAGCAGTTG 4098
RESULT 2
AR401379 LOCUS 4481 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 20 from patent US 6623927.
ACCESSION AR401379
VERSION AR401379.1 GI:40148692
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4481)
AUTHORS Brahmachari,S.K., Choudhry,S., Mukerji,M. and Jain,S.
TITLE Method of detection of allelic variants of SCA2 gene
JOURNAL Patent: US 6623927-A 20 23-SEP-2003;
Council of Scientific and Industrial Research; New Delhi;
INX;
FEATURES
source Location/Qualifiers
1..4481
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 2.26e-97 Length: 4481
Pred. No.: 6961.00 Matches: 1312
Score:
```

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-802-228-2 (1-1312) x AR401379 (1-4481)			
QY	1	MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe	20
DB	163	ATCGCGTCAGCGCGCGAGCTCTCGGAGTCCGCGGTGGCCACCGAGTCTCGCGGCTTC	222
QY	21	AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaAlaArgSerGlyArg	40
DB	223	GCGCGAGCAGGTGGCGCGCGGTGGCGCTCGCTCCAGCGCGCGCGCGCGGCGG	282
QY	41	GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro	60
DB	283	GCGCGCGGTGGCGCGCGCGCGGACGATATCCCTCCGCGCGCGCTCCCGCGCGCGCC	342
QY	61	GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn	80
DB	343	GGCGCGCGCTCCCTCCCGCGAGAGCTCGCTCCCTCCCGCTCAGACTGTTTTGGTAGCAAC	402
QY	81	GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro	100
DB	403	GGCAAGCGCGCGCGCGGTTCGGCGCGCGCTCCCGCGCGCTCCTTGGTCTCGCGCGGCT	462
QY	101	ProArgProPheValValValLeuLeuProLeuAlaSerProGlyValAlaProProAlaAla	120
DB	463	CCCCCGCGCTTGTGTGTGTCTTCTCCCTCCGCGCGCGCGCGCGCGCTCCCGCGCGG	522
QY	121	ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu	140
DB	523	CCAAACCGCGCTCCCGCTCGCGCGCGGTGCTCCCGCGCGCTCCCGCGCTCTCTTG	582
QY	141	AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr	160
DB	583	GCGCGCGCGCTCCCGCGCTGTCGCGCGCGCGGTGCGAGCGCGGTATGGCGCGCTCAC	642
QY	161	MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	180
DB	643	ATGTCTCTGAAGCCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	702
QY	181	GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	200
DB	703	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	762
QY	201	SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerSerSerSer	220
DB	763	AGCGGCGCTTCTAGCGTTCGCGCGCGCGCGCGCTTCCGCGCTCTCGGTCTCTCG	822
QY	221	SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly	240
DB	823	TCTCTGGCCACCGCTCCCTCTCTGCTGGTGGCGGACCTCCGCGCGCGGAGCGCGG	882
QY	241	LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly	260
DB	883	CTGGGCAGAGGTGGAACACAGTAACAAAGACTGCGCTCAGTCTAGATTTCTTTTATGGA	942
QY	261	IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu	280
DB	943	ATCTATGCAATATAGGATGTTTATATATCTTACATCAGTTGTTGGCTCCCAATGTAA	1002
QY	281	ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	300
DB	1003	GTACAAAGTGAAGAGGTATATGAAGAGGTTTTTAAACCTTACAGTCCGAAAGTGT	1062
QY	301	AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg	320
DB	1063	GATTTGGTACTTGATGCCGCATCAGAAAGTACAGATTCAGTTCGGGCGCGAAACGT	1122
QY	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys	340
DB	1123	GAAGAAATAATGAGAGATATTTTGTTCAAATGTTTTCAGACTTTGTGTGGTACAGTTTAAA	1182
QY	341	AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys	360
DB	1183	GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTTAAA	1242
QY	361	ValAsnGlyGluHisLysGlyLeuAspLeuGluProTrpAspAlaGlyGluLeuThrAla	380
DB	1243	GTGAATGCGGAACACAAAGAGAGACTCTGGAGCCCTGGATGCGAGTGAACCTCACAGCC	1302
QY	381	AsnGluGluLeuLeuAlaLeuLeuAsnAspValSerAsnGlyTrpAspProAsnAspMet	400
DB	1303	AATGAGGAACCTTGAAGCTTTTCGAAATGACGTATCTAATGATGGATCCCAATGATATG	1362
QY	401	PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer	420
DB	1363	TTTTCATATATGAAGAAATATATGTTGTTAGTGTCTACGTATGATACGAGTTTATCTTTC	1422
QY	421	TyrThrValProLeuLeuArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla	440
DB	1423	TATACAGTCCCTTAGAAGAGATAACTCAGAAGAAATTTTAAACGGGAAGCAAGGCA	1482
QY	441	AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu	460
DB	1483	AAACCAAGTTAGCAGAGAAATTTGAGTCAAGTCCCGCAGTACAAAGCTCGAGTGGCCCTG	1542
QY	461	AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg	480
DB	1543	AATGATATAGAGTGAAGAAATATACACAGAGTTTCAGAGAAATTTCCAGTGAACGT	1602
QY	481	GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg	500
DB	1603	GAGGGGACACAGATAAACACACTAGGGAAAAATAATATATCTCTCGACAAAGAAATAGA	1662
QY	501	GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer	520
DB	1663	GAAAGTCATATCTCGGGAAAGTGGGAGACAGAAATTCACGCGCTATGCGCCAGCCTTGA	1722
QY	521	GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer	540
DB	1723	GGCTCCATGCCATCAAGATCCACTTCTCACACTTTCAGATTTTCAACCCGAAATTTCTG	1782
QY	541	AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer	560
DB	1783	GACCAAGAGTAGTTAATGAGAGGTTCCTCGGCGCATCGCTTGGCCCATCTCTCTCTCT	1842
QY	561	ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro	580
DB	1843	CGCCACCTTCTCGCTACCGTCCAGTCCCAACTCTCTTCCAGCTCGGGCAGCACCCCT	1902
QY	581	ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSer	600
DB	1903	ACACGGCGCGCTCCAGCGCGCTCCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCT	1962
QY	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
DB	1963	CATGGTCTCCAGCT	2022
QY	621	ArgMetSerProLysAlaGlnArgHisProArgAsn	

```
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleProThrGluAlaValAla 720
DB 2263 CCCAGTGGGCCAGTCTTCTCCAGCTGGTATATTATTCACACTGGAAGCTGTTC 2322
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGCCATTCCAGCTGCATCTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnArgGlnAsnSerProAlaGly 760
DB 2383 CCTTCTAGTGAGGCTAAAGATTCCAGCTTCAAGATCAGAGGAGAGAACTCTCTCTG 2442
QY 761 AsnLysGluAsnLysProAsnGluThrSerProSerPheSerIlyshAlaGluAsnLys 780
DB 2443 AATAAAGAAAATATTAAACCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysPheLys 800
DB 2503 GGTATATACCAAGTGTGTTCTGNACATAGAAAACAGATTGATGATTTAAAGAAATTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 AATGATTTTAGGTTACAGCAAGTCTCTCTCTGAATCTATGATCAACTATAACAAA 2622
QY 821 AsnArgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2623 AATAGAGAGGGAGAAAATCAAGAGATTGATCAAAAGACAAAATTTGAACAAAGTGCTAAG 2682
QY 841 AspSerPheIleGluAsnSerSerAsnGlyThrSerGlySerSerLysProAsnSer 860
DB 2683 GATTCCTTTCAATGAATAAGCAGCAGCAACTGTACAGTGAGCAGCAAGCGCAATAGC 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
DB 2743 CCAGCATTTCCCTTCAATCTAGTACAGGAGCACAAGAGGGGAGCTGAGGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
DB 2803 TCCAAAGGGGTTCCAGACTTCCAGCCAGCATGTAAACAGAGAAAGACGATAGGAAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
DB 2863 AAGAAAGAGCAGCTGAGCAAGTTAGGAATCAACATTAATCCCAATGCAAGGAGGTTC 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
DB 2923 AACCCAGTTCCTTCTCTCAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCTCAA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValThrGlnPro 960
DB 2983 GCACAACTAGCCCATCTATGTGGGTCTATCAACAGCCAACTCCAGTTTATATCTAGCCT 3042
QY 961 valCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
DB 3043 GTTTGTTTGGCAAAATATGATGTATCCAGTCCCAAGTCCAGGCGGCGTGCACCTTTA 3102
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
DB 3103 TACCCATATCTATGAGCCCATGCGCATGGAATCAAGCCCAAGACATATAGAGCAGTACCA 3162
QY 1001 AsnMetProGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
DB 3163 AATATGCCCCAACAGCGGCAAGACCCAGCATCATCAGAGTGCCATGATGACCCAGCGTCA 3222
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
DB 3223 GCAGCGGGCCACCGATGTCAGCCACCCCAAGCTTACTCCAGCAATATGTTGCTTAC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
DB 3283 AGTCTCAGCAGTTTCCAAATAGCCCCCTTGTTCAGCATGTGCCACATATCATCTCAG 3342
```

```
QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
DB 3343 CATCTCATGTCTTAGTCTCTGTAATACAGGTAATGCTAGTAATGATGGCACCACAACA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
DB 3403 CACGCCCAAGCTGGTTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATCAGCAGACG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
DB 3463 CATGGCATGTATGTCATGTCCCAAAATTACCATACAAAGGAGACAGGCCCTCTTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
DB 3523 TTTGCCATTTCCACGGGCTCCCTGCTCAGCAGTATGGCACCTTAACGCTACCTTGCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
DB 3583 CCACATATCTCCACACCTCAGCTTTCAGTACCTCCCTCAGCAGCAGCAAGCCAAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
DB 3643 GGTGGAAGTCTATCTCTGCACTCCCTGCTCAGCAGCAGTACAGCAGCAGCCCGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTACAGCAGTATTTACCACGGGGCTTGGCCT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
DB 3763 CCACCTCTCCATGACACCTGCTCCAAACAGCAGTCCGACAGATAATAGTTTCCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
DB 3823 CAACAGACTGTCTTTACGATCCATCTCTCAGTTCAGCCGCGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3883 CACATGGGCCCATCTACCTCAGCTCATGTACAGTACAGGAATGGTTCCTTCTCATCCA 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaLeu 1280
DB 3943 GCCCATGCGCCCAATGATGCTAATGACGACACAGCAGCCCGGCTCCCGAGCCGCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
DB 4003 GCTCAAGTGCACTACAGCCCATTCAGTCTCGACAAACAGCGCATTTCCCTTATATGACG 4062
QY 1301 HisProSerValGlnAlaHisHisGlnGlnLeu 1312
DB 4063 CACCTTTCAGTCAAGCCCAACCCACCAACAGCAGTTG 4098
RESULT 3
LOCUS AR447280 4481 bp mRNA linear PAT 20-FEB-2004
DEFINITION Sequence 2 from patent US 6673535.
ACCESSION AR447280
VERSION AR447280.1 GI:42675575
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4481)
AUTHORS Pulst,S.M.
TITLE Methods of detecting spinocerebellar Ataxia-2 nucleic acids
JOURNAL Patent: US 6673535-A 2 06-JAN-2004;
FEATURES Location/Qualifiers
source 1. 4481
/organism="unknown"
/mol_type="mRNA"
ORIGIN
```


Alignment Scores:

Pred. No.: 2,26e-97 Length: 4481
Score: 6361.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-802-228-2 (1-1312) x AR447280 (1-4481)

Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 163 ATGCGCTCAGCGCGCGCGAGCTCTCGGAGTCCGCGGTGGCCACCGAGTCTCGCGCTTC 222
Qy 21 AlaAlaAlaArgTrrProGlyTrrArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db 223 GCCCGAGCCAGGTGGCGCGCGGTGGCGTCCGTCAGCGCGCGCGCGCGGAGCGGCGG 282
Qy 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 283 GCGCGCGGTGGCGCGCGCGCGGACGATATCCCTCCGCGCGCGCGCGCGCGCGCGCG 342
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db 343 GCG 402
Qy 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100
Db 403 GCGAAACGCGCGCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 462
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAla 120
Db 463 CCG 522
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 523 CCAACCG 582
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 583 GCG 642
Qy 161 MetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 643 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
Qy 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
Db 763 AGCGCGCTTCTAGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 822
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
Db 823 TCCTCGGCGCAGCGCTCCCTCTCGGTGGTTCGCGCGCGCGCGCGCGCGCGCGCGCG 882
Qy 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
Db 883 CTGGGCGAGGTTCGAAACAGTAAACAGATGCGCTCAGTCTACGATTTCTTTTATGGA 942
Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
Db 943 ATCTATGCAAAATATGAGGATGGTTTCATATATTCATCAGTTGTTGGCTCCAAATGTGA 1002
Qy 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db 1003 GTACAAGTGAAGAAATGGAAGGTATATATGAAGAGTATTTTAAACCTTACAGTCCGAAGTGT 1062
Qy 301 AspLeuValLeuAspAlaAlaHisGlyLysSerThrGluSerSerSerGlyProLysArg 320
Db 1063 GATTTGGTACTTGATGCGCGCATGAGAAAGTACAGAAATCCAGTTCGCGGCGCGAAACGT 1122

Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
Db 1123 GAAGAAATAATGAGAGATTTTGTTCAAATGTTTCAGACTTGTGTGTCAGATTTAAA 1182
Qy 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db 1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTCAA 1242
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTrrAspAlaGlyGluLeuThrAla 380
Db 1243 GTGAATGCGGAACACAAAGAGAGAGCTTGAGCGCTGGGATGCGAGTGAATCAGACCC 1302
Qy 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrrAspProAsnAspMet 400
Db 1303 AATGAGGAATCTGAGGCTTTGGAAAAATGACGTATCTAATGATGGATGCCCAATGATG 1362
Qy 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
Db 1363 TTTGATATAATGAAGAAAAATATGGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db 1423 TATACAGTGCCTTTAGAAAGAGATAAATCAGAAAGATTTTTTAAACCGGAAGCAAGGCA 1482
Qy 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db 1483 AACCACTTAGCAGAAAGAAATTCAGTCAAGTCCCGCAGTACAAAGCTCGAGTGCCTGAA 1542
Qy 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db 1543 AATGATGATAGAGGTGAGGAAGAAAAATACACAGCAGTTCAGAGAAATTTCCAGTGAACGT 1602
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500
Db 1603 GAGGGGCACAGCATAAACACTTAGGGAATAATAATATATCTCTCCGACAAAGAAATAGA 1662
Qy 501 GluValIleSerTrrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1663 GAAGTCATATCTCGGGAAAGTGGGAGACAGAAATTCACCGGTATGGCGCAGCTGGATCG 1722
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1723 GGTCTCAATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATCTTGTTCA 1782
Qy 541 AspGlnArgValValAsnGlyValProTrrProSerProCysProSerProSerSer 560
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCCTTGCCCATCTCTCTCTCT 1842
Qy 561 ArgProProSerArgTrrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1843 CGCCACCTTCTCGCTACCACTACAGTCAAGTCCCAACTCTCTTCCACCTCGGGCAGCCACCT 1902
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db 1903 ACACGGCGCGCTTCAGCGCGCGCTTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1963 CATGGTCTTCAGCTCTCTGCTCTACTATGCTTAACGATGTCTTTCAGAAAGGCGCTCCA 2022
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Db 2023 AGGATGTCCTCCAAAGGCGCCAGCAGCATCTTCGAAATCAGAGATTTCTGTGGGAGGGGT 2082
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
Db 2083 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACACCCAGTGAAGCAGTACTCTCT 2142
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTrrPrrSerSerValValSerGlyValPro 680
Db 2143 CCAGTAGCAGGACAGTCCCTCGGGGGGAAAGTGTGTATCATCAGTGTGTGAGTGGGTCCA 2202

QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
 DB 2203 AGATTATCCCTAAACACTATAGACCAGGTCTCCAGACAGAAAGTATTGGAAATACC 2262
 QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
 DB 2263 CCAGTGGGCCAGTTCTTGCTCTCCCAAGCTGGTATTATTCCTCAACTGAACTGTGGC 2322
 QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
 DB 2323 ATGCCTATTCCAGCTGCATCTCTACGCTGCTAGTCTCTGCATCGAACAGAGCTGTATCC 2382
 QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAlaGlnArgGlnAsnSerProAlaGly 760
 DB 2383 CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAGATCAGAGGACAGAACTCTCTCGAGG 2442
 QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
 DB 2443 AATAAGAAATATTAAACCCCAATGAACATCACCTAGTCTTCTCAAACTGAAACAA 2502
 QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspLeuLysLysPheLys 800
 DB 2503 GGTATATCACCACTGTTCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2562
 QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
 DB 2563 AATGATTTTAGGTTACAGGCAAGTCTACTTCTGAATCTATGATCAACTACTTAACAA 2622
 QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
 DB 2623 AATAGAGAGGAGAAATCAAGAGATTGATCAAGACAAATTTGACCAAGTGTCTAAG 2682
 QY 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
 DB 2683 GATTCCTTCAATGAATAGCAGCACTGTACCACTGGCAGCAGCAAGCCGATAGC 2742
 QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
 DB 2743 CCCAGCATTTCCCTCTCAATCTAGTAACACGAGCACAAGAGGGGAGCTTGAGGTCACT 2802
 QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
 DB 2803 TCCCAAGGGGTTCCAGACTTCCAGCCAGCATGTAAACAGAGAAACCATAGGAAGAG 2862
 QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
 DB 2863 AAGAAAGACGACGCTGAGCAAGTTAGGAATCAACATTGATCCCAATGCAAGGAGTTC 2922
 QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
 DB 2923 AACCCAGTTCTCTCTCAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCCTCAA 2982
 QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValThrThrGlnPro 960
 DB 2983 GCACAACTAGCCCATCTATGGTGGGTCATCAACAGCCAACTCCAGTTATATCTCAGCCT 3042
 QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
 DB 3043 GTTGTGTTGCACCAATATGATGATCCAGTCCAGTGAGCCAGCGGTGCAACCTTTA 3102
 QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
 DB 3103 TACCCCAATACCTATGACGCCCATGCCAGTGAATCAAGCCAAACATATAGAGCAGTACCA 3162
 QY 1001 AsnMetProGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
 DB 3163 AATATGCCCCAACAGCGGCAAGACCAACATCATCAGAGTGCCATGATGACCCAGCGTCA 3222
 QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
 DB 3223 GCAGCGGGCCCGATTGAGCCACCCAGCTTACTCCAGGCATATGTTGCTTAC 3282
 QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060

DB 3283 AGTCTTCAGCAGTTCCTCCCAATCAGCCCTTGTTTCAGCATGTGCCACATTATCAGTCTCAG 3342
 QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
 DB 3343 CATCTCTATGCTCTAGTCTCTGTAATACAGGGTAATGCTAGAACTGATGGCACCACCAACA 3402
 QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
 DB 3403 CACGCCACCGCTGGTGTATGATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGAG 3462
 QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
 DB 3463 CATGGATGTATGCTATGCTCCCAATTTACATACAACAGGAGCAAGCCCTTCTTCTTAC 3522
 QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
 DB 3523 TTTGCCATTTCCACGGGCTCCCTTCTCAGCAGTATGCGACCTTAACGCTACCTTGCAC 3582
 QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
 DB 3583 CCACATATCTCCACACCTCAGCTTTCAGTACCCCTTGGACAGCAGCAAGCCCAACAT 3642
 QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
 DB 3643 GGTGAAGTCACTCTGACCCAGTCTGTTTCAGCACCATCAGCAGCAGCCGCCAGGCT 3702
 QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
 DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTTCAGCCATTTTACCACGGGGCTTCGCCAACT 3762
 QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
 DB 3763 CCACCTCTCCATGACACCTGCTCCAAACAGCAGTCCGCCACAGAATAGTTTCCAGCAGCA 3822
 QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
 DB 3823 CAACAGACTGCTTTTACGATCCATCTCTTCTCAGTTCAGCCGGGTATACCAACCCACCC 3882
 QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
 DB 3883 CACATGGCCACGTCACCTCAGGCTCATGTACAGTCAGGAATGTTCTTCTTCTCATCCAAC 3942
 QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaLeu 1280
 DB 3943 GCCCATGCCCCAATGATGCTAATGACGACACAGCCACCCGGGGTCCCGAGCCGCCCTC 4002
 QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
 DB 4003 GGTCAAAGTGCACTACAGCCCATTCAGTCTCGAACACAGCGCATTTCCCTATATGACG 4062
 QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
 DB 4063 CACCCTTCAGTACAGCCCAACCCACCAACAGCAGTTCG 4098

RESULT 4

AR632098 4481 bp mRNA linear PAT 14-FEB-2005
 LOCUS Sequence 2 from patent US 6844431.
 DEFINITION AR632098
 ACCESSION AR632098
 VERSION AR632098.1 GI:59773824
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 AUTHORS 1 (bases 1 to 4481)
 TITLE Nucleic acid encoding spinocerebellar ataxia-2 and products related thereto
 JOURNAL Patent: US 6844431-A 2 18-JAN-2005;
 FEATURES Cedars-Sinai Medical Center; Los Angeles, CA
 source Location/Qualifiers
 1. .4481

/organism="unknown"									
/mol_type="mRNA"									
ORIGIN									
Alignment Scores:									
Pred. No.:	2.26e-97	Length:	4481						
Score:	6961.00	Matches:	1312						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	6	Gaps:	0						
US-10-802-228-2 (1-1312) x AR632098 (1-4481)									
Qy	1	MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe	20	301	AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg	320	Qy	301	AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg
Db	163	ATGCGCTCAGCGCGCGAGCTCCTCGAGTCCCGGTCGCCAGTTCGCGCGTTC	222	1063	GATTTGGTACTTGTATGCGGCACATGAGAAAGTAGTACAGAATCCAGTTCGGGGCCGAAACGT	1122	Db	1063	GATTTGGTACTTGTATGCGGCACATGAGAAAGTAGTACAGAATCCAGTTCGGGGCCGAAACGT
Qy	21	AlaAlaAlaArgTTPProGlyTTPArgSerIeuGlnArgProAlaArgSerGlyArg	40	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys	340	Qy	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys
Db	223	GCGGAGCCAGGTGGCGCGGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGG	282	1123	GAAGAAATAATGGAGAGTATTTTGTTCAAATGTTCAGACTTTGTGTGTACAGTTTAAA	1182	Db	1123	GAAGAAATAATGGAGAGTATTTTGTTCAAATGTTCAGACTTTGTGTGTACAGTTTAAA
Qy	41	GlyGlyGlyValAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro	60	341	AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys	360	Qy	341	AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys
Db	283	GGCGGCGTGGCG	342	1183	GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA	1242	Db	1183	GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA
Qy	61	GlyGlyGlyValAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro	80	361	ValAsnGlyGluHisLysGluLysAspLeuGluProTTPAspAlaGlyGluLeuThrAla	380	Qy	361	ValAsnGlyGluHisLysGluLysAspLeuGluProTTPAspAlaGlyGluLeuThrAla
Db	283	GGCGGCGTGGCG	342	1243	GTGAATGGCGAACAACAAGAGAGGACCTGGAGCCCTGGGATGCAGGTGAACACTCACAGCC	1302	Db	1243	GTGAATGGCGAACAACAAGAGAGGACCTGGAGCCCTGGGATGCAGGTGAACACTCACAGCC
Qy	61	GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn	80	381	AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTTPAspProAsnAspMet	400	Qy	381	AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTTPAspProAsnAspMet
Db	343	GGCCCCCTCCTCCCGGCAGAGCTCGCCTCCCTCCCGCTCAGACTGTTTGTAGCAAC	402	1303	AATGAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGATG	1362	Db	1303	AATGAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGATG
Qy	81	GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro	100	401	PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer	420	Qy	401	PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer
Db	403	GGCAACGGCGCGCGGCTTTCGGCGCGGCTCCCGCGCGCTCCTTGGTCTCGGGCGGCGCT	462	1363	TTTCGATATAATGAAGAAAAATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG	1422	Db	1363	TTTCGATATAATGAAGAAAAATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG
Qy	101	ProArgProPheValValValLeuLeuProLeuAlaSerProGlyValaProProAlaAla	120	421	TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla	440	Qy	421	TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla
Db	463	CCCCGCCCTTCGTGTGTCTTCTCCCTTCGCCAGCGCGGCGCGCTCCCGCGCGCGG	522	1423	TATACAGTGCCCTTAGAAAGAGATAAATCAGAAGAAATTTTAAACGGGAAGCAAGGCA	1482	Db	1423	TATACAGTGCCCTTAGAAAGAGATAAATCAGAAGAAATTTTAAACGGGAAGCAAGGCA
Qy	121	ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu	140	441	AsnGlnIleuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu	460	Qy	441	AsnGlnIleuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu
Db	523	CCAAACCGCGCTCCCGCTCGCGCGCGTGGTTCGCCCGCGCGTCCCGCGCTCCTTG	582	1483	AACAGTGTAGCAGAAAGAAATTTAGTCAAGTGGCCAGTCAAAAGCTCGAGTGGCCCTGGAA	1542	Db	1483	AACAGTGTAGCAGAAAGAAATTTAGTCAAGTGGCCAGTCAAAAGCTCGAGTGGCCCTGGAA
Qy	141	AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr	160	461	AsnAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg	480	Qy	461	AsnAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg
Db	583	GCGCGCGCGCTCCCGGCTGTCCCGCGCGCGGTGCGAGCGGTATGGCGCGCTCACC	642	1543	AATGATGATAGGAGTGGAGAAAGAAATACACAGCAGTTTCAGAGAAATTCACAGTGAACGT	1602	Db	1543	AATGATGATAGGAGTGGAGAAAGAAATACACAGCAGTTTCAGAGAAATTCACAGTGAACGT
Qy	161	MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	180	481	GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg	500	Qy	481	GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg
Db	643	ATGTCGCTGAACCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	702	1603	GAGGGCACAGCATAAACACACTAGGAAAAATAATATATATCTCTCTGGACAAAGAAATAGA	1662	Db	1603	GAGGGCACAGCATAAACACACTAGGAAAAATAATATATATCTCTCTGGACAAAGAAATAGA
Qy	181	GlnGlnGlnGlnGlnGlnProProProAlaAlaAlaAsnValArgLysProGlyGly	200	501	GluValIleSerTTPGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer	520	Qy	501	GluValIleSerTTPGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer
Db	703	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	762	1663	GAAGTCATATCTCGGGAAGTGGGAGACAGAAATTCACCGCGTATGGGCGACGCTGGATCG	1722	Db	1663	GAAGTCATATCTCGGGAAGTGGGAGACAGAAATTCACCGCGTATGGGCGACGCTGGATCG
Qy	201	SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer	220	521	GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer	540	Qy	521	GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer
Db	763	AGCGGCTTCTAGCGTCCCGCGCGCGCGCTTCGCCGTTCCTCGTCTCCTCGTCTCG	822	1723	GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCCGAATTCCTGGTTCA	1782	Db	1723	GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCCGAATTCCTGGTTCA
Qy	221	SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly	240	541	AspGlnArgValValAsnGlyGlyValProTTPProSerProCysProSerProSerSer	560	Qy	541	AspGlnArgValValAsnGlyGlyValProTTPProSerProCysProSerProSerSer
Db	823	TCCTCGGCCACGGCTCCCTCCTCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	882	1783	GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGGCATCGCCTTGCCCATCTCCTCTCTCT	1842	Db	1783	GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGGCATCGCCTTGCCCATCTCCTCTCTCT
Qy	241	LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly	260	561	ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro	580	Qy	561	ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro
Db	883	CTGGGACAGAGTCAAGAAACAGTAACAGACTGCTCAGTCTACGATTTCTTTTATGGA	942	1843	CGCCACCTTCTCGCTACCAAGTCCCAACTCTCTTCACCTCGGCGGACCCACCTCT	1902	Db	1843	CGCCACCTTCTCGCTACCAAGTCCCAACTCTCTTCACCTCGGCGGACCCACCTCT
Qy	261	IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu	280	581	ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla	600	Qy	581	ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla
Db	943	ATCTATGCAAAATATGAGATGGTTTCATATACATCATGATGTTGGTTCGCTCCAAATGTGAA	1002	1903	ACACGGCGCCCTCCAGGCCCTCCGCGGCATTCAGACCCCTCTCACCCCTCTGCT	1962	Db	1903	ACACGGCGCCCTCCAGGCCCTCCGCGGCATTCAGACCCCTCTCACCCCTCTGCT
Qy	281	ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	300	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620	Qy	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro
Db	1003	GTACAAGTGAAATGGAGGTATATATGAAAGAGTTTAAAACTTACAGTCCGAGTGT	1062	1963	CATGGTCTCCAGCTCCTGTCTCTACTATGCTTAAAGCATGTCTTCAGAAAGGGGCTCCA	2022	Db	1963	CATGGTCTCCAGCTCCTGTCTCTACTATGCTTAAAGCATGTCTTCAGAAAGGGGCTCCA
Qy	300	ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	320	621	ArgMetSerProLysAlaGlnArgHisProArgHisArgValSerAlaGlyArgGly	640	Qy	621	ArgMetSerProLysAlaGlnArgHisProArgHisArgValSerAlaGlyArgGly
Db	1062	GTACAAGTGAAATGGAGGTATATATGAAAGAGTTTAAAACTTACAGTCCGAGTGT	1122	2023	AGGATGTCCCAAGGGGCCGAGCATCCTCGAAATCACAGAGTTTCTGCTGGGAGGGGT	2082	Db	2023	AGGATGTCCCAAGGGGCCGAGCATCCTCGAAATCACAGAGTTTCTGCTGGGAGGGGT
Qy	320	ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	340	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro	660	Qy	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro
Db	1122	GTACAAGTGAAATGGAGGTATATATGAAAGAGTTTAAAACTTACAGTCCGAGTGT	1182	2083	TCCATATCCAGTGGCTTAGAATTTGTATCCCAACCCAGTGAAGAGCTACTCTCT	2142	Db	2083	TCCATATCCAGTGGCTTAGAATTTGTATCCCAACCCAGTGAAGAGCTACTCTCT

QY 661 ProValAlaArgThrSerProSerGlyThrTrpSerSerValValSerGlyValPro 680
DB 2143 CCAGTAGCAAGACCAAGTCCCTCGGGGGAACGTGGTCAATCAGTGGGTTCAC 2202
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
DB 2203 AGATTATCCCTTAAAACTCATAGACCAGGTCTCCAGACAGAAAGATATTGGAATACC 2262
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
DB 2263 CCCAGTGGCCAGTCTTCTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGC 2322
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGCCTATTCCAGCTGCATCTCTACGCTGTAGTCTCTGTCATCGAACAGAGCTGTATAC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2383 CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCGAAGTCTCTCGAGGG 2442
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2443 AATAAGAAATATTAAACCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
DB 2503 GGTATATCACCACTGTTCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 AATGATTTTAGGTTAGCAGCAAGTCTTACTTCTGAATCTATGGATCAACTACTTAAACAAA 2622
QY 821 AsnArgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2623 AATAGAGGGAGAAATCAAGAGATTGATCAAGAGCAAAATTCGACNAGTGTCTAG 2682
QY 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
DB 2683 GATTCCTTCAATGAAATAGCAGCAGCAACTGTATCCAGTGGCAGCAGCAAGCCGAATAGC 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
DB 2743 CCAGCATTTCCCTTCAATCTAGTATTAACACGAGCAAGAGGGGACCTGAGGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnLysAspLysGluGlu 900
DB 2803 TCCCAAGGGTTCCAGCTTCAGCCCGCATGTATTAACAGAGAAAGCATAGGAAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
DB 2863 AAGAAAGACGACGTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAGGAGTTTC 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
DB 2923 AACCCACGTTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTCACCTCGGCCCTCA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyThrGlnPro 960
DB 2983 GCACAACCTAGCCCATCTATGGTGGTTCATCAACAGCCCACTCAAGTTTATCTCAGCCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyProValProValSerProGlyValGlnProLeu 980
DB 3043 GTTTGTTTTCACCAATATGATGTATCCAGTCCAGTGAGCCAGCGGTGCAACCTTTA 3102
QY 981 TyProIleProMetThrProMetProValAsnGlnAlaLysThrTyArgAlaValPro 1000
DB 3103 TACCCAAATACCTATGACGCCCATGCCAGTGAATCAAGCCCAAGACATATAGACAGTACCA 3162
QY 1001 AsnMetProGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
DB 3163 AATATGCCCAACAGCGGCAAGACCAGCATCATCAGATGCCATGATGACCCAGCGTCA 3222
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyThrGlnTyValAlaTyThr 1040

DB 3223 GCAGCGGGCCCAACGATTTCAGAGCCACCCACAGCTTACTCCACCAATATGTTCCCTTAC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyGlnSerGln 1060
DB 3283 AGTCTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAG 3342
QY 1061 HisProHisValTySerSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
DB 3343 CATCTCATGTCTATAGTCTCTGTAATACAGGTAATGCTAGAAATGATGACCAACCAACA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyGlyValHisGlnGlnThr 1100
DB 3403 CACGCCCAAGCTGGTGTATCTTCTTCCAGCACTCAGTACGGGCTCATGAGCAGACG 3462
QY 1101 HisAlaMetTyAlaCysProLysLeuProTyAsnLysGluThrSerProSerPheTyThr 1120
DB 3463 CATGCGATGTATGTCATGTCCTCCAAATTTACCATACAACAGGAGACAAGCCCTTCTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyThrAlaHisProAsnAlaThrLeuHis 1140
DB 3523 TTTTGCATTTTCCAGCGGCTCCCTTGTCTCAGCAGTATGCGCACCTAACGCTACCTGCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
DB 3583 CCACATATCTCCACACCTCAGCTTTCAGTACCTCCCTCAGTACGAGCAGCAAGCAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180
DB 3643 GGTGGAAGTATCTCTGACCCAGTCTCTGTTTACGACCATCAGCAGCCGCGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyHisAlaGlyLeuAlaProThr 1200
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTTCAGCCTTTTACACGCGGGCTTGCCTCACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
DB 3763 CCACCTCTCATGACACCTGCTCCCAACAGCAGTCCGACAGCAAGTATGTTCCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyThrAsnProPro 1240
DB 3823 CAACAGCTGTCTTTACGATCTCTTCTCAGCTTTCAGCCGGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3883 CACATGGCCCACTACCTCAGGCTCATGTACAGTACAGTACAGGAATGTTCTTCTCATCACT 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
DB 3943 GCCCATGCGCCCAATGATGCTAATGACGACACAGCCACCCGCGGTCCCCAGGCCGCCCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyMetThr 1300
DB 4003 GCTCAAGTGCATCAGCCCATTCAGTCTCGACAAACAGCGCATTTCCCTTATATGACG 4062
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
DB 4063 CACCTCTGATGATAGTACAGCAAGCCCAACCAACAGCAGTTG 4098

RESULT 5

AX392465 4481 bp DNA linear PAT 23-MAR-2002
LOCUS Sequence 1 from Patent WO0216417.
DEFINITION AX392465
ACCESSION AX392465
VERSION AX392465.1 GI:19700764
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Pulst,S.M. and Huynh,D.P.

Db 1603 GAGGGGCAAGCATAAACACTAGGAGAAATAATAATATATTCCTCCGGCAAGAAATAAGA 1662
QY 501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1663 GAAGTCATATCTCTGGGAAGTGGAGACAGAATTACCGCGTATGGCCAGCCTGGATCG 1722
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACATTCAGATTTCACGCCGAATTCCTGGTTCA 1782
QY 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCTTGCCCATCTCTCTCCCT 1842
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1843 CGCCCACTTCTCGCTACCGTCCAGTCCCACTCTCTTCCACTCGGCGAGCCACCCCT 1902
QY 581 ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla 600
Db 1903 ACAGGCGCCCTCCAGGCGCCCTCGCGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1963 CATGGTCTCCAGCTCCTGTCTACTATGCGCTTAAACGCGATGCTTTCAAGAGGCGCTCCA 2022
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Db 2023 AGATGTCTCCCAAGGCCAGCGACATCTCGAATCAAGATTTCTGCTGGAGGGT 2082
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
Db 2083 TCCATATCCAGTGGCTAGAAATTGTATCCCAACCCAGTGAAGCAGTACTCTCT 2142
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTyrProSerSerValValSerGlyValPro 680
Db 2143 CCAGTAGCAAGGACCAAGTCCCTCGGGGGGAACTGGTCAATCAGTGTGAGTGGGGTTCCA 2202
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db 2203 AGATTTATCCCTTAAACTATAGCCAGGTCTCCAGACAGAACAGTATGGAAATACC 2262
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
Db 2263 CCCAGTGGGCGAGTCTCTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGC 2322
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
Db 2323 ATGCCTATTCCAGTGCATCTCTACGCTGTAGTCTCTGCATCGAACAGAGCTGTACC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
Db 2383 CCTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATCAGAGGCGAGAACTCTCTCTGAGGG 2442
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
Db 2443 AATAAGAAATAATTAACCAATGAACATCATACCTAGCTTCTCMAAAGCTGAACACAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysPheLys 800
Db 2503 GGTATATACCAAGTTGTTCTGAACATAGAAAAAGATTCATGATTTAAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db 2563 AATGATTTAGTTTACAGCCAAAGTTTACTTCTGAATCTATGGATCAACTACTATAACAA 2622
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
Db 2623 AATAGAGGGGAGAAAAAATCAAGAGATTGATCAAGACAAAAATTTGAACCAAGTGTAAAG 2682
QY 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860

Db 2683 GATTCTTTTCATTGAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCCAATAGC 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db 2743 CCCAGCATTTCCCTTCAATACTTAGTAACACGGAGCAACAAGAGGGGACCTTGAGGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900
Db 2803 TCCCAGGGGTTTACAGCTTCCAGCCAGCATGTAAACAAGAGAAAGACGATAAAGGAAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 2863 AAGAAAGACGAGCTGACCAAGTTAGGAAATCAACATTTGAATCCCAATGCCAAAGAGATT 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2923 AACCCACGCTTCTCTCAGCCAAAGGCTTCTACTACCCCAACTTTCACCTCGGCGCTCAA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db 2983 GCACAACCTAGCCCATCTATGGTGGGTCAATCAACAGCCCACTCCAGTTTATACTCAGCCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db 3043 GTTTCGTTTGCACCAATATGATGATCCAGTCCAGTGAGCCAGGCGTGCAACCTTTA 3102
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 3103 TACCCAATACCTTATGACGCCCATGCGAGTGAATCAAGCCCAAGACATATATAGACGAGTACCA 3162
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db 3163 AATATGCCCAACAGCGGCAAGCAGCATCATCAGAGTCCCATGATGACCCACCGCTCA 3222
QY 1021 AlaAlaGlyProProIleAlaAlaThrProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 3223 GCAGCGGGCCCAACGATTCAGCCACCCACAGCTTACTCCACGCAATATGTTGCCTAC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 3283 AGTCCTCAGCAGTTCCTCCAAATCAGGCCCTTGTTCAGCATGTGCCATTTATCAGTCTCAG 3342
QY 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCTCTCATGTCTATAGTCTGTAAATACAGGGTAATGCTAGAAATGATGACCAACCAACA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Db 3403 CAGCCCGACCTGGTGTAGTATCTTCTTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3463 CATGCGATGTATGATGTCTCCAAATTTACCATACAAAGAGAGCAAGCCCTCTCTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTGCGATTTCCACGGGCTCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCTGCGAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
Db 3583 CCACATATCCACACCTCAGCCTTCCAGTACCCCACTGGACAGCAGCAAGCAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAla 1180
Db 3643 GGTGGAAGTCACTCTGCACCCAGTCTCTGTTTTCAGCAGCATCAGCACCGGCGCCAGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3703 CTCATCTGGCCAGTCCACAGCAGCAGTCAAGCATTTTACCACCGGGGCTTGGCCCAACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3763 CCACCTTCCATGACCTGCTCTCAACACGAGTCCGCCAGAAATAGTTTCCCAGCAGCA 3822

QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
DB 883 CTGGGAGAGGTGGAACAGTAAACAAAGAGCTGCCTCAGTCTACGATTTCTTTTGATGA 942
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
DB 943 ATCTATGCAATATAGAGTGGTTTCATATACATCATCATGTTGCTCCAAATGTGAA 1002
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB 1003 GTACAAAGTGAATAATGAGGTATATATGAAGAGTGTAAAACTTACAGTCCGAAGTGT 1062
QY 301 AspLeuValLeuAspAlaHisGluLysSerThrGluSerSerGlyPProLysArg 320
DB 1063 GATTTGGTACTTGATCCGCACATGAGAAAGTACAGATCCAGTTCGGGGCGGAAACGT 1122
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
DB 1123 GAAGAATAATGAGAGTATTTTGTTCAAATGTTTCAGACTTTGTGTGTTACAGTTTAA 1182
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB 1183 GATATGGAATCCAGTTATGCAAAAGAGAGTCTTTTACTGACTCTGCTATCAGTGTAAA 1242
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380
DB 1243 GTGAATGGCGAAACACAAAGAGAGGACCTGGAGCCCTGGGATGCAGGTGAACCTCACAGCC 1302
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
DB 1303 AATGAGGAACCTTGAGGCTTTGGAAATGACGTATCTAATGGATGGGATCCCAATGATATG 1362
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB 1363 TTTGATATATAGAGAAATTTATGGTGTAGTCTACGTATGATAGCAGTTTATCTTCG 1422
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB 1423 TATACAGTGCCTTAGAGAGAGATAACTCAGAGAAATTTTAAACGGGAAGCAAGGGCA 1482
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1483 AACCAAGTACAGAGAAATTTAGTCAAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTGGAA 1542
QY 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB 1543 AATGATGATAGGAGTGGAGAGAAATAACACAGCAGTTCAGAGAAATTCACGTGAACGT 1602
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
DB 1603 GAGGGGCACAGCATAAACACTAGGGAAATAAATATATATCTCTCTGGACAAAGAAATAGA 1662
QY 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB 1663 GAAGTCAATATCTCTGGGAAAGTGGAGACAGAAATTCACCGGTATGGCCAGCCTGGATCG 1722
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1723 GGCTCCATGCGATCAAGATCACTTCTACACTTCAGATTTCAACCCGAATTCGTGTCA 1782
QY 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
DB 1783 GACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCTGCCCATCTCTCTCTCTCT 1842
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
DB 1843 CGCCACCTTCTCGCTACCAAGTCAAGTCCCAACTCTCTCCACTCTCGGGCAGCCACCT 1902
QY 581 ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla 600
DB 1903 ACAGGGCGCCCTCCAGGCCCTCCGGGCGCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620

DB 1963 CATGTTCTCCAGCTCTCTGTCTCTACTATGCTAAACGATGTCTTTCAGAGGGCTCCA 2022
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
DB 2023 AGGATGTCCCCAAAGGCCAGCGACATCTCTGAAATCACAGAGTTTCTGCTGGAGGGT 2082
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
DB 2083 TCCATATCAGTGGCTAGAAATTTGTATCCCAACACCACCAGTGAACGACTACTCTCT 2142
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680
DB 2143 CCAGTAGCAGACAGTCCCTCGGGGGAACTCGTCATCAGTGGTCACTGGGGTTCCA 2202
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
DB 2203 AGATTATCCCTTAAACCTCATAGACCAGTCTCCAGACAGACAGTATTGGAAATACC 2262
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaAla 720
DB 2263 CCAGTGGGCCAGTTCTTGTCTTCTCCCAAGCTGGTATTATTCCAACTGAAAGCTGTGGC 2322
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGCTATTTCAGCTGCATCTCTACGCTGCTAGTCTGCATCGAACAGAGCTGTACC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2383 CCTTCTAGTAGGCTTAAAGATTCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGCAGG 2442
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2443 AATAAAGAAATATTAAACCCAATGAACCATCACCTAGTCTCTCAAAAGCTGAAACAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
DB 2503 GGTATATCACCAAGTGTGTTCTGAACATAGAAACAGATTGATTTAAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 AATGATTTTAGGTACAGCCAAAGTCTACTTCTGAATCTATGATCAACTACTATAACAA 2622
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2623 AATAGAGAGGAGAAATCAAGAGATTGTGATCAAAAGACAAATTTGAACCAAGTGTAA 2682
QY 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
DB 2683 GATTTCTTCAATGAAATAGCAGCAACTGTACCAGTGGCAGCAGCAGCCGAATAGC 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
DB 2743 CCAGCATTTCCCTTCAATCTTAGTAACCGAGCACAAAGAGGGGACCTGAGGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
DB 2803 TCCCAAGGGGTTCCAGACTTCCAGCCCGCATGTAAACAAGAGAAACGATAAGGAAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
DB 2863 AAGAAAGAGCGAGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAGGAGTTC 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
DB 2923 AACCCAGTTCCTTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACCTCGGCCTCAA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
DB 2983 GCACAACCTAGCCCATCTATGGTGGGTCTCAACAGCCAACCTCAGTTTATCTACGCCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980

Db 3043 GTTTGTTTGCACCAATATGATGATATCCAGTCCAGGAGCCCGGTGCAACCTTTA 3102
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 3103 TACCCAATACCTTATGAGCGCCCATGCGAGTGAATCAAGCAACATATAGAGCATCA 3162
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db 3163 AATATGCCCCAAACAGCGGCAAGACCATCATCATCAGAGTGCCATGATGACCCAGCTCA 3222
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 3223 GCAGGGGGCCCAACGATTCAGCCACCCACAGCTTACTCCACCATATATGTTGCCATC 3282
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 3283 AGTCCTCAGCAGTCCCAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAG 3342
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCCTCATGTCTATAGTCTCTGTAATACAGGGTAATGCTAGAAATGATGGCACCAACA 3402
Qy 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Db 3403 CAGGCCAGCCTGGTTTGTATGATCTTCTCAGCAACTCAGTACGGGGCTCATGAGCAGCG 3462
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3463 CATCGATGTATGATGCTGCCAAATACCATACACAGGAGACAGCCCTCTTCTTCTAC 3522
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTGGCAATTCACGGGCTCCCTTGTCTCAGCAGTATGGCACCTTAACGCTACCCCTGCAC 3582
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db 3583 CCACATACTCCACACCTCAGCCTTCTCAGTACCCCACTGGACAGCAGAAAGCCAAACAT 3642
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
Db 3643 GGTGAAGTCACTCTCTGCACCCAGTCTGTTCAGACCATCAGCACAGCGCCGCGAGCT 3702
Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCATTTTACCAACGCGGGCTTTCGCCAACT 3762
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3763 CACCCCTCATGACACCTGCTCCAAACAGCAGTCCGACAGAAAGTATTTCCAGCAGCA 3822
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3823 CAAACAGCTGTCTTTACGATCATCTCTCTCAGTTCAGCGCGGTATACCAACCCACCC 3882
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3883 CACATGGGCCACGTCCTCAGCCTCATGTATACAGTCCAGAAATGGTTCCTTCTCATCCAAC 3942
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3943 GCCCATGGCCCAATGATGCTAATAGCAGACAGCCACCGCGGGTCCCGAGCGCCCTC 4002
Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrAlaHisPheProTyrMetThr 1300
Db 4003 GCTCAAGTGCACTACAGCCCAATTCAGTCTCGACAAACAGCGCAATTTCCCTATATGAG 4062
Qy 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db 4063 CACCCTTCAGTACAAGCCCAACCAACAGCAGTTG 4098

RESULT 7
CQ727533

. LOCUS CQ727533 4479 bp DNA linear PAT 03-FEB-2004

DEFINITION Sequence 13467 from Patent WO02068579.
ACCESSION CQ727533
VERSION CQ727533.1 GI:42293778
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
human exons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 13467 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1. .4479
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.65e-96 Length: 4479
Score: 6890.00 Matches: 1308
Percent Similarity: 99.77% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 3
Query Match: 98.98% Indels: 2
DB: 6 Gaps: 0
US-10-802-228-2 (1-1312) x CQ727533 (1-4479)
Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 163 ATGGCTCAGCGCGCGAGCTCTCGAGTCCCGGGTGGCCACCGAGTCTCGCGG-TTC 221
Qy 21 AlaAlaAlaArgTgTgProGlyTgTgArgSerLeuGlnArgProAlaArgArgSerGlyArg 40
Db 222 GCGCAGCAGGTGCGCGGGTGGCGTCTCGCTCCAGCGCGCGCGCGCGCGCGCGG 281
Qy 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 282 GCGCGCGTGGCGGGCGCGCGGACCGGTATCCCTCCGCGCCCTTCCTCCCGCGCGCC 341
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db 342 GCGCGCGCTCCCTCCGCGCAGCTCGCTCCCTCCGCTCAGACTGTTTGTAGCAAC 401
Qy 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100
Db 402 GCGAACGCGCGCGCGCTTTTCGCGCCGCTCCCGCGGCTCCTTGGTCTCGCGCGGCT 461
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
Db 462 CCGCGCGCTTCGTCT 521
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 522 CCAACCGCGCTCCCGCTCGCGCGCGCGCTCCCGCGCGCTCTCTCTCTCTCTCTCT 581
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 582 GCGCGCGCGCTCCCGCTGTCTCCCGCGCGCGCGCGCGCGCGGTGTATGGGCCCTCAC 641
Qy 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 642 ATGTCGCTGAAGCCCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 701
Qy 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 702 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 761

QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerValSerSer 220
DB 762 AGCGGCTTTAGAGCTGCGCGCGCGGCT-TGCGCGCTCGCTCGCTCGCTCG 820
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
DB 821 TCTCTGGCCACGGCTCCCTCTCGGTGGTGGCGGACCTCCGGCGGGAGGCGCGC 880
QY 241 LeuGlyArgGlyArgAsnSerAsnGlyLeuProGlnSerThrIleSerPheAspGly 260
DB 881 CTGGGCGAGGTCGAACAGTAGAACAGACTGCCTCAGCTACGATTTCTTTGATGGA 940
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
DB 941 ATCTATGCAAAATATGAGTAGGTTCATATACATTACATCAGTTGTTGGCTCCAAATGTGAA 1000
QY 281 ValGlnValIleAsnGlyGlyIleTyrGluGlyValPheIleThrTyrSerProLysCys 300
DB 1001 GTACAAAGTGAATAATGGAGTATATATGAGGAGTTTTTAAACCTTACAGTCCGAAGTGT 1060
QY 301 AspLeuValIleAspAlaAlaHisGluLysSerThrGluSerSerGlyProLysArg 320
DB 1061 GATTTGGTACTTGATCCGCACATGAGAAAGTACAGATCCAGTTCCGGGCGGAAACCT 1120
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
DB 1121 GAAGAAATAATGAGAGTATTTGTTCAATGTTACAGACTTGTGTTGAGTATTA 1180
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB 1181 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTCTATCAGTGTCTAAA 1240
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380
DB 1241 GTGAATGGCGAACACAAAGAGAGGACCTGGAGCCCTGGGATCGAGTGAATCAGAGCC 1300
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
DB 1301 AATGAGGAACTTGAGGCTTTGGAAATGACGTATCTAATGGATGGATCCCAATGATATG 1360
QY 401 PheArgTyrAsnGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB 1361 TTTTCGATATAATGAAGAAATTTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTCG 1420
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB 1421 TATACAGTGCCTTAGAAGAGATTAATCAGAGAAATTTTAAACGGGAGCAGGGCA 1480
QY 441 AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1481 AACCAAGTTAGCAGAGAAATTTAGTCAAGTCCAGGTACAAAGCTCGAGTGGCCCTGGAA 1540
QY 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB 1541 AATGATGATAGGAGTGAGGAGAGAAATATACACAGCAGTTTCAGAGAAATTTCCAGTGAACGT 1600
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
DB 1601 GAGGGCACAGCATAAACCTAGGAAATAATATATATTTCTCTGTGACAAAGAAATAGA 1560
QY 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB 1661 GAAGTCATATCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGCCAGCTGGATCG 1720
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1721 GGCTTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATTCGTGTTCA 1780
QY 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560
DB 1781 GACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCTTGGCCCATCTCTCTCTCT 1840
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580

DB 1841 CGCCACACTTCTCGCTACAGTCCCACTCTCTTCCACTCGGGCAGCCACCCCT 1900
QY 581 ThrArgProProSerArgProSerArgProSerArgProProSerHisProSerAla 600
DB 1901 ACACGGCGCGCTCCAGGCCCCCTCGGGCCATCCAGACCCCGCTCTCACCCCTTGCT 1960
QY 601 HisGlySerProAlaProValIleSerThrMetProLysArgMetSerSerGluGlyProPro 620
DB 1961 CATGGTTCTCCAGCTCTCTCTACTATGCTAAACGATGTCTTTCAGAGGGGCTCCA 2020
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
DB 2021 AGGATGTCCCCAAGGCCAGCGACATCTCGAAATCACAGAGTTTCTCTCGGAGGGT 2080
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
DB 2081 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACCCAGTGAAGCAGCTACTCCT 2140
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTyrSerSerValIleSerGlyValPro 680
DB 2141 CCAGTAGCAAGGACCACTCCCTCGGGGGAACGTGGTTCATCAGTGGTGGGTGCCA 2200
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyValAsnThr 700
DB 2201 AGATTATCCCCTAATAACTCATAGACCCAGGTCTCCAGACAGAACAGTATTCGAATACC 2260
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
DB 2261 CCCAGTGGCCAGTTCTTGTCTTCTCCCAAGCTGTATTTATTTCCAACTGAACTGTGGC 2320
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2321 ATGCTATTCCAGCTGCATCTCTACGCTGTAGTCTCTGCATCGAACAGAGCTGTATCC 2380
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2381 CCTTCTAGTAGGCTTAAAGATTCCAGGTTCAAGATCAGAGGCGAGAACTCTCTCGAGGG 2440
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2441 AATAAAGAAATATTTAAACCCCAATGAACATCACCTAGTCTTCTCAAAAGCTGAAACAAA 2500
QY 781 GlyIleSerProValIleSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
DB 2501 GGTATATCACCAAGTTTCTGAAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2560
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2561 AATGATTTTAGGTTACAGCAAGTTCTACTTCTGAAATCTATGGATCAACTACTATAACAAA 2620
QY 821 AsnArgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2621 AATAGAGAGGAGAAAAATCAAGAGATTTGATCAAGACAAAAATTTGAAACCAAGTGTAG 2680
QY 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
DB 2681 GATTTCTTCAATGAAATAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAATAGC 2740
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
DB 2741 CCAGCATTTCCCTTCAATACTTAGTAACCGGAGCACAAAGAGGGGACCTCGAGGTCACT 2800
QY 881 SerGlnGlyValGlnThrSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900
DB 2801 TCCCAAGGGGTTACAGCTTCCAGCCCGCATGTAAACAAGAGAAACCATTAAGGAAGAG 2860
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
DB 2861 AGAAAGACGAGCTGAGCAAGTTAGGAAATCAACATTCGAATCCCAATGCAAGGAGTTTC 2920
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940

Db 2921 AACCCAGCTTCCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAA 2980

Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960

Db 2981 GCACAACCTAGCCATCTATGGTGGGTCAACAAGCCAACTCCAGTTTATACTCAGCCT 3040

Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980

Db 3041 GTTTGTTTTGCACCAATATGATGATCCAGTCCCAAGTGGCCCAAGGGGTGCAACCTTTA 3100

Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000

Db 3101 TACCCAAATACCTATGAGCCCAATGCAAGTGAATCAAGCAAGACATATAGAGCAGTACA 3160

Qy 1001 AsnMetProGlnGlnArgAlaAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020

Db 3161 AATATGCCCAACACGGGCAAGACCAGCATCATCAGAGTGCCATGATGCACCCAGCGTCA 3220

Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040

Db 3221 GCAGCGGGCCCAACCATTTGCAGCCACCCACAGCTTACTCCACGCAATATGTTGCCTAC 3280

Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060

Db 3281 AGTCCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAG 3340

Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080

Db 3341 CATCCTCATGCTATAGTCTGTATACAGGGTATGCTAGATGATGCCACCAACA 3400

Qy 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100

Db 3401 CACGCGCCAGCCTGGTTAGTATCTCTTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3460

Qy 1101 HisAlaMetTyrAlaCysProIleLeuProTyrAsnLysGluThrSerProSerPheTyr 1120

Db 3461 CATGGCATGTATGCAATGCCCCAAATACCATACAAGGAGACAAAGCCCTCTCTTCTAC 3520

Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140

Db 3521 TTTGCCATTCACGGGCTCCCTGCTCAGCAGTATGCCACCTTAACGCTACCTCGAC 3580

Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160

Db 3581 CCACATACTCCACACCTCAGCCTTTCAGTACCCCTCAGTACGACAGCAGCAAGCCAACT 3640

Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180

Db 3641 GGTGAAGTATCTCTGCACCCAGTCTGTTTCAGCACCATCAGCACCGGCCCGCCAGGCT 3700

Qy 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200

Db 3701 CTCATCTGGCCAGTCCACAGCAGCAGTACGATTTACCACGCGGGCTTGGGCCAAT 3760

Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220

Db 3761 CCACCCCTCCATGACACCTGCTTCCCAACGCGAGTGCACAGAAATAGTTTCCAGCAGCA 3820

Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240

Db 3821 CAACAGACTGTCTTTACGATCCATCTCTCATCGTTACGCGGGGTATACCAACCCACCC 3880

Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260

Db 3881 CACATGGCCCACTACCTCAGGCTCATGTACAGTACAGATGGTTCTCTCATCCAAT 3940

Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280

Db 3941 GCCCATGGCCCAATGATGCTAATGACGACACAGCCACCGCGGGTCCCGAGCGCCCTC 4000

Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300

Db 4001 GCTCAAGTGCTATAGCCCCATTTCAGTCTCGACACAGCGCATTTCCCTTATATGAGC 4060

Qy 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312

Db 4061 CACCTTTTCAGTACAAAGCCCAACCAACAGCAGTGTG 4096

RESULT 8

HSDANSCA2

LOCUS HSDANSCA2 4163 bp mRNA linear PRI 18-APR-2005

DEFINITION H.sapiens mRNA for SCA2 protein.

ACCESSION Y08262

VERSION Y08262.1 GI:1770389

KEYWORDS SCA2 gene.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS Imbert,G., Saudou,F., Yvert,G., Devys,D., Trotter,Y., Garnier,J.M., Weber,C., Mandel,J.L., Cancel,G., Abbas,N., Durr,A., Didierjean,O., Stevanin,G., Agid,Y. and Brice,A.

TITLE Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats

JOURNAL Nat. Genet. 14 (3), 285-291 (1996)

PUBMED 8896557

REFERENCE 2 (bases 1 to 4163)

AUTHORS Imbert,G.

TITLE Direct Submission

JOURNAL Submitted (20-SEP-1996) G. Imbert, I.G.B.M.C., Departement Of Genetics, B.P. 163, 67404 Illkirch Cedex, FRANCE

FEATURES

source

1..4163

/organism="Homo sapiens"

/mol_type="mRNA"

/isolate="DNA patient"

/db_xref="taxon:9606"

/cell_line="lymphoblastoid"

/clone_lib="DAN"

/dev_stage="adult"

<1..4163

/gene="SCA2"

<1..2747

/gene="SCA2"

/codon_start=3

/protein_id="CAA69589.1"

/db_xref="GI:1770390"

/db_xref="COA:099700"

/db_xref="UniProt/Swiss-Prot:Q99700"

/translation="GNGGGAFRPGSRRLGLGGPRPFVVVLLPLASPGAPAAPTRA SPLGARASPRSGVSLRAPCAPCPACEPVYGLTMSLKPQQQQQQQQQQQQQQQ QQQQPPAAANVRKPGSGLLASPAAPSPSSSVSSSSATAPSVVAATSGGCRPG LGRGRNKGLPQSTISFDGIYANRMVHILTSVVGSKCEVQVKNGGIYEGVFTYSP KCDVLDAAEHKESTESSGPKREIMESILFKCSDFVVVOFKMDSYAKEDPTDSA ISAKVGHKEKLEPMDAGELTANESLEALENDVSNMGDFNDMFRINEENYGVVSTY DSSLSSTYVPLERDNEBEFLKEARANQAEIESAQYKARVALENDDRSEEEKYTA VQRNSREGHSINTRENKTIIPGQRNREVIISWGSRQNSPRMQQSGMSRSTSH TSDPNNSGSDQRVVNGKVPWPCPSPPSRPSYQSGNSLPRAATPTRPSPRP SPSPSPSPSHAGSPAPVSTMPKRMSEGGPPRMSPKAQRHPRNHRYSAGRGSISSGL EFVSHNPSEAAATPPARTSPSGTWSVSGVPLRLSPKTHRPSPRSPQNSIGNTPSPG VLAGSPQAGLIPTEAVAMPPIPAASPTPASPSNRAVTPSESAKSRJQQRONSAPGNK ENIKPNSTSPFSKAENKGISPVVSEHRKQIDLLKFKNDPRLQPSSTSEMDOLNK NREGEKRDITKDKIIPESAKDSFIENSSNCTSGSKSPNSPSISPSISUNTEHRGPE VTSQGVTSPPACKQEKDKKEKDAAEQVRKSTLNPNAKEFNPRSPSQPKPTPTS PRPQAQSPSPSMVGHQFTPVVYTPQVCPAPNMVPPVPSVPGVQYQICPNSGKTSIRVP "

ORIGIN

Alignment Scores:

Pred. No.: 1,34e-88 Length: 4163

Score: 6377.00 Matches: 1211

Percent Similarity: 98.30% Conservative: 0

Best Local Similarity: 98.30% Mismatches: 1

Query Match: 91.61% Indels: 21


```
Db      2103 GGTATATCACCAGTTGTTTCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2162
QY      801  AenAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db      2163 AATGATTTTAGGTTACAGCCAAGTTCTACTTCTGAATCTATGGATCAACTCTAAACAAA 2222
QY      821  AenArgGluGlyGlyLysSerArgAspLeuLysAspLysLysLeuProSerLysLys 840
Db      2223 AATAGAGAGGGAGAAAATCAAGAGATTGTGATCAAAAGACAAAATTTGAACCAAGTCTAAG 2282
QY      841  AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Db      2283 GATTTCTTTCATTTGAANAATAGCAGCGCAACTGTACAGTGGCAGCAGCAGCCGGAATAGC 2342
QY      861  ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db      2343 CCCAGCAVTTCCCTTCAATACTTAGTAACCGGAGCACAAGAGGGGACCTGAGGTCACT 2402
QY      881  SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGlyLysAspLysGluGlu 900
Db      2403 TCCCAAGGGGTTTCAGACTTCCAGCCCGCAGCATGTAACCAAGAGAAAGACGATAAGGAAGAG 2462
QY      901  LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db      2463 AAGAAAGAGCGCAGCTGAGCAAGTTAGGAAATCAACATTTGAATCCCAATGCCAAGAGATTC 2522
QY      921  AenProArgSerPheSerGlnProLysProSerThrProThrSerProArgProGln 940
Db      2523 AACCCACGTTCTCTCTCAGCCAAAGCCTTCTACTCCCCCAACTTCACTCGGCCCTCAA 2582
QY      941  AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db      2583 GCACAACTTAGCCCCATCTATGGTGGGTCTATCAACAGCCCAACTCCCACTTTATACTCAGCCT 2642
QY      961  ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db      2643 GTTTGTTTGGACCAATATGATGATCCAGTCCAGTCCAGCCCGGGGTGCATATA---- 2698
QY      981  TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db      2699 -----CCCA 2701
QY      1001 AenMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db      2702 AATATGCCCCAACACAGCGCAAGACCATCATCATCAGAGTGCCTATGATCACCACGCGTCA 2761
QY      1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db      2762 GCAGCGGGCCCAACGATTGACGCCACCCACCGCTTACTCCACGCAATATGTTGCCCTAC 2821
QY      1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db      2822 AGTCCTCAGCAGTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAG 2881
QY      1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db      2882 CATCTCATGTCTATAGTCTCTGTAATACAGGGTAATGTCTAGAAATGATGGCACCAACA 2941
QY      1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Db      2942 CACGCCCAAGCTGGTTTATGATCTCTTTCAGCACTCAGTACGGGGCTCATGAGCAGACG 3001
QY      1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db      3002 CATGCGATGATGATGATGCCAAATACCATACAAAGGAGACAAAGCCCTCTTCTTCTAC 3061
QY      1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db      3062 TTTGTCATTTCCACGGGCTCCCTTGTCTCAGCAGTATGGCCACCTTAACGCTACCTGAC 3121
QY      1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db      3122 CCACATATCTCCACCCCTCAGCTACCCCTCAGTACCCCTCAGCAGCAGCAAGCAACCAT 3181
```

```
QY      1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
Db      3182 GGTGGAAGTCATCTCTGCACCAGTCTCTGTTTTCAGCACCATCAGCAGCCGCCAGGCT 3241
QY      1181 LeuHisIleuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db      3242 CTCCATCTGGCCAGTCCACAGCAGCAGTACGCCATTTACCAACGCGGGCTTTCGCCAAT 3301
QY      1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db      3302 CCACCTCTCATGACACCTGCTCCACACGCGCTCCGACAGATAGTTTCCAGCAGCA 3361
QY      1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db      3362 CAACAGACTGTCTTTAGCATCCATCTCTCAGCTTACGCGCGGTATACCAACCCACCC 3421
QY      1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db      3422 CACATGGCCCCAGTACCTCAGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAAT 3481
QY      1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db      3482 GCCCATGGCCCAATGATGCTAATGACGACACAGCACCAGCGGTCCCGAGCGCCCTC 3541
QY      1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
Db      3542 GCTCAAAAGTGCACTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTTATATGAG 3601
QY      1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312
Db      3602 CACCTTTTCAGTACAAGGCCACCACCAACAGCAGTTG 3637

RESULT 9
LOCUS   A62706 4200 bp DNA linear PAT 12-MAR-1998
DEFINITION Sequence 7 from Patent WO9717445.
ACCESSION A62706
VERSION A62706.1 GI:3716590
KEYWORDS
SOURCE  unidentified
        unclassified sequences.
REFERENCE
  1. Tora, L., Lutz, Y., Trottier, Y., Mandel and Jean-Louis.
    METHOD FOR TREATING NEURODEGENERATIVE DISEASES USING A 1C2 ANTIBODY
    OR A FRAGMENT OR DERIVATIVE THEREOF, AND CORRESPONDING
    PHARMACEUTICAL COMPOSITIONS
    JOURNAL Patent: WO 9717445-A 7 15-MAY-1997;
    COMMENT CENTRE NAT RECH SCIENT (FR)
    FEATURES Other publication FR 2741088 19970516.
            Location/Qualifiers
            source 1..4200
                  /organism="unidentified"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:32644"
                  /clone="DAN1"

ORIGIN

Alignment Scores: 1.35e-88 Length: 4200
Pred. No.: 6377.00 Matches: 1211
Score: 98.30% Conservative: 0
Percent Similarity: 98.30% Mismatches: 1
Best Local Similarity: 91.61% Indels: 21
Query Match: 6 Gaps: 1
DB:

US-10-802-228-2 (1-1312) x A62706 (1-4200)

QY      81  GlyAsnGlyGlyGlyAlaPheArgProGlySerArgLeuLeuLeuGlyLeuGlyPro 100
Db      3  GGCAACGGCGCGCGCGTTCGGCGCCCGGCTCCCGCGGGCTCTTGGTCTCGCGGGGCT 62
```

QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
DB 63 CCGCGCCCTTCGTCTGCTCTCTCCCTCGCCAGCCCGCGCGCCCTCGCGCGCG 122
QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB 123 CCACCCGCGCCCTCCCGCTCGCGCGCCGTGGTCCCGCGCGGTTCGGCGTCTCTTG 182
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB 183 GCGCGCCCGCTCCCGCTGTCGCCCGCGCGGTGCGAGCGGTGTATGGCGCCCTCAC 242
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 243 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 302
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 303 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 362
QY 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
DB 363 AGCGGCTTCTAGCGTCGCGCGCGCGCGCGCTTCGCGCTCTCGTCTCTCTCG 422
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
DB 423 TCCTCGGCACCGCTCCCTCGGTGTCGCGCGACCTCCGCGCGCGCGAGCCCGCG 482
QY 241 LeuGlyArgGlyArgAnSerAnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
DB 483 CTGGCAGAGGTCCGAACAGTAACAAAGGAGTCCCTCAGTCTACGATTTCTTTGATGGA 542
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
DB 543 ATCTATGCAATATAGGATGGTTTATATATCTATCATCATCATGTTTGGCTCCAAATGTGAA 602
QY 281 ValGlnValLysAnGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB 603 GTACAGTGAAATAGGAGTATATATGAAGGAGTTTTTAAACCTTTACAGTCCGAAAGTGT 662
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
DB 663 GATTTGGTACTTGATCGGCACATGAAAGATGACAGAAATCCAGTTGGGGCCGAAACGT 722
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
DB 723 GAAGAATAATGAGAGATATTTGTTCAATGTTTACAGCTTTGTTGTGTACAGTTTAA 782
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB 783 GATATGACCTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTAA 842
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla 380
DB 843 GTGAATGGCGAAACAAAGAGAGACCTGGAGCCCTGGGATGCGAGTGAACCTCACAGCC 902
QY 381 AsnGluIleLeuGluAlaLeuGluAsnAspValSerAnGlyTrpAspProAsnAspMet 400
DB 903 AATGAGGAATCTGAGGCTTTGGAAAAATGAGCTATCTAATGGATGGGATCCCAATGATATG 962
QY 401 PheArgTyrAnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB 963 TTTTCGATATATGAGAAATATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1022
QY 421 TyrThrValProLeuGluArgAspAnSerGluGluPheLysArgGluAlaArgAla 440
DB 1023 TATACAGTGCCTTAGAAAGAGATAACTCAAGAAATTTTAAACCGGAAGCAAGGCGCA 1082
QY 441 AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1083 AACCAAGTTAGCAGAAAGAAATTTAGTCAAGTGCCAGTACAAAGCTCAGTGGCGCCCTGGAA 1142
QY 461 AsnAspArgSerGluGluGluLysTyrThrAlaValGlnArgAnSerSerGluArg 480

DB 1143 AATGATGATAGGAGTCAGGAAGAAAAATACAGCAGTTTCAGAGAAATTCAGTCAACGT 1202
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAnArg 500
DB 1203 GAGGGGCACAGCATAAACACATAGGAAAAATAATATATTCCTCTCGGACAAAGAAATAGA 1262
QY 501 GluValIleSerTrpGlySerGlyArgGlnAnSerProArgMetGlyGlnProGlySer 520
DB 1263 GAAGTCATATCTCGGGGAAGTGGGAGAGAGAATTCACCGCGTATGGCCAGGCTTGATCG 1322
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1323 GGCTCATGCCATCAAGATCCACTTCTCACACTTTCAGATTTCACCCGAAATTCGTGTTCA 1382
QY 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
DB 1383 GACCAAGAGTAGTTAATGGAGGTGTCCTGGCCATCGCTTGCCTTGCCTCTCTCTCTCT 1442
QY 561 ArgProProSerArgTyrGlnSerGlyProAnSerLeuProProArgAlaAlaThrPro 580
DB 1443 CGCCCAACCTTCTCGCTACCCAGTCAGGTCCCAACTCTCTTCCACCTCGGCGACCCCTT 1502
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
DB 1503 ACACGGCGCCCTCGAGGCCCTTCGCGGCCATCCAGACCCCGTCTCACCCCTCTGCT 1562
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
DB 1563 CATGTTCTCCAGCTCTCTCTACTATGCTTAAACGATGCTTTCAGAGGGCTTCCA 1622
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAnHisArgValSerAlaGlyArgGly 640
DB 1623 AGGATGTCGCCAAAGGCCCGCAGACATCTCGAAATACACAGAGTTTCTCTCGGAGGGGT 1682
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProSerGluAlaAlaThrPro 660
DB 1683 TCCATATCCAGTGGGCTAGAAATTTGATCCCAACACCCAGTGAAGCAGTCTACTCCT 1742
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680
DB 1743 CCAGTAGCAAGGACCAAGTCCCTCGGGGGAACGTGTCATCAGTGTCTAGTGGGTTCCA 1802
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAnSerIleGlyAsnThr 700
DB 1803 AGATTATCCCTAAAACTCATAGACCAGGTCTCCAGACAGAACTATTTGAAATATACC 1862
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleProThrGluAlaValAla 720
DB 1863 CCCAGTGGGCCAGTTCTTCTCTCCCAAGCTGGTATTTATTCCAACTGAAGCTGTGGCC 1922
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 1923 ATGCTCTATTCAGCTGCATCTCTCGCTCTAGTCTCTGATCGAACAGAGCTGTATACC 1982
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAnSerProAlaGly 760
DB 1983 CCTTCTAGTGAGGCTAAAGATTTCCAGCTTCAAGATCAGAGGCGAGAACTCTCTCGAGGG 2042
QY 761 AsnLysGluAnLysProAnGluThrSerProSerPheSerLysAlaGluAnLys 780
DB 2043 AATAAGAAAAATATTAAGCCCAATGAACATCACTAGCTTCTCAAAAGCTGAAAAACAA 2102
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysPheLys 800
DB 2103 GGTATATCACCAAGTTGTTCTGAACATAGAAAAACAGATTTGATGATTTTAAAGAAATTTAAG 2162
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAnLys 820
DB 2163 AATGATTTAGGTATACAGCAAGTTCTACTTCTGAATCTATGATCACTACTTAAACAA 2222
QY 821 AsnArgGluGlyGluLysSerArgAspLeuLysAspLysIleGluProSerAlaLys 840

Db 2223 AATAGAGGGGAGAAAAATCAAGAGATTGTATCAAAAGACAAAATTTGAACCAAGTCTTAAG 2282

Qy 841 AspSerPheIleGluAsnSerSerAenCysThrSerGlySerSerIysProAsnSer 860

Db 2283 GATTCTTTTCATTGAAAATAGCAGCAGCAACTGTACCAAGTGGCAGCAGCAAGCCGAATAGC 2342

Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880

Db 2343 CCCAGCATTTCCCTTCAATACTTGTAAACCGSAGCAACAGAGGGGACCTGAGGTCACT 2402

Qy 881 SerGlnGlyValGlnThrSerSerProLaCysLysGlnGluLysAspAspLysGluGlu 900

Db 2403 TCCCAAGGGTTTCAGACTTCAGCCCGCAGCATGTAACAGAGAAAGACGTAAGAGAGAG 2462

Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920

Db 2463 AAGAAAGACGCAGCTGAGCAAGTTAGGAAATCAACATTGAATCCCCTCAATGCAAGAGGTTTC 2522

Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940

Db 2523 AACCCACGTTCTTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCTCAA 2582

Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960

Db 2583 GCACAACTTAGCCCATCTATGGTGGGTCTAACAAGCCCACTCCAGTTTATCTACTCAGCCT 2642

Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980

Db 2643 GTTTGTTTGGCACCATAATATGATGATCCAGTCCAGTGAGGCCAGGCGTCAATA 2698

Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000

Db 2699 -----CCA 2701

Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020

Db 2702 AATATGCCCAACAGCGGACAGCAGCATCATCAGAGTGCCATGATGACCACCGGTCA 2761

Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040

Db 2762 GCAGCGGGCCACCGATTGTCAGCCACCCACCGACTTACTCCAGCAATATGTTGCCTAC 2821

Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060

Db 2822 AGTCCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATATCAGTCTCAG 2881

Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaAtqMetMetAlaProProThr 1080

Db 2882 CATCTCATGTTATAGTCTGTAATACAGGTAATGCTAGTAATGATGACACCAACA 2941

Qy 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100

Db 2942 CAGCGCCAGCCTGGTTTAGTATCTTCTTCAGCAACTCAGTACGGGCTCATGAGCAGAGC 3001

Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120

Db 3002 CATGCGATGTATGATGTGCCAAATTACCATAACAAGAGAGCAAGCCCTTCTTTCTAC 3061

Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnTyrAlaHisProAsnAlaThrLeuHis 1140

Db 3062 TTTGGCATTTTCAGCGGCTCCCTTGTCTCAGCAGTATGGCACCCCTAACGCTACCCCTGCAC 3121

Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160

Db 3122 CCACATACTCCACACCTCAGCTTCAGTACCCCACTGGACAGCAGCAAGCCAAAT 3181

Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180

Db 3182 GGTGGAAGTCATCTCGACCCAGTCTGTTACGACCATCAGCACAGGCCGCCAGGCT 3241

Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200

Db 3242 CTCCTATCTGGCAGTCCACAGCAGCAGTCAAGCATTTTACCACGCGGGGCTTCGCGCAACT 3301

Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220

Db 3302 CCACCTCCATGACACCTGCCCTCCAAACACGAGTCCGACAGAATAGTTTCCAGCAGCA 3361

Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240

Db 3362 CAACAGACTGTCTTTAGCATCCATCTCTCACGTTACGCCGGGTATACCAACCCACCC 3421

Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260

Db 3422 CACATGGCCACGTAACCTCAGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAAC 3481

Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280

Db 3482 GCCCATGGCCCAATGATCTAATGACGACACACAGCCAGCCGCGGTCCTCCAGCCGCCCTC 3541

Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300

Db 3542 GCTCAAGTGCACACTACAGCCCATTCAGTCTCGAACAACAGCGCATTTCCCTATATGACG 3601

Qy 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312

Db 3602 CACCTTCAGTACAAGCCCAACCAACACAGCAGTTG 3637

RESULT 10

AF041472

LOCUS

DEFINITION Mus musculus ataxin-2 (SCA2) mRNA, complete cds.

ACCESSION AF041472

VERSION AF041472.1 GI:3005019

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 4225)

AUTHORS Nechiporuk, T., Huynh, D.P., Figueroa, K., Sabha, S., Nechiporuk, A. and Pulst, S.M.

TITLE The mouse SCA2 gene: cDNA sequence, alternative splicing and protein expression

JOURNAL Hum. Mol. Genet. 7 (8), 1301-1309 (1998)

PUBMED 9668173

REFERENCE 2 (bases 1 to 4225)

AUTHORS Nechiporuk, T.T., Figueroa, K., Sabha, S., Nechiporuk, A.V. and Pulst, S.M.

TITLE Direct Submission

JOURNAL Submitted (07-JAN-1998) Medicine/Neurology, Cedars-Sinai Medical Center, 8700 Beverly Blvd., Los Angeles, CA 90048, USA

FEATURES

Location/Qualifiers

1..4225

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/chromosome="12"

/map="12q23.1"

1..4225

/gene="SCA2"

27..3884

/gene="SCA2"

/codon_start=1

/product="ataxin-2"

/protein_id="AAC09275.1"

/db_xref="GI:3005020"

/translation="MRSSTAAVQRPAAQDPEPRPAGWAARRSLPRTARRGGRGAVAYPSAGPPRGGAPGRSPSPSCDCFSNGHSGASRPSRLILGVCPPRPFFVVVLLALAPAATPRACPPGVRASPPRSGSSARPAAGCPACPEPVGPTLMSLKPPQPPAPATKRPGGGLSSPGAAPASAAVTSASVPAAPAVASSAAAGRPGILGRGNS SKGLPPTISFDGIYANVRMVHILTSVVGSKCEVQKNGGIYEGVFTKYSKCDLVLDAHSEKSTESSGPKREIMESVLFKCSDFVVFQKDTDSVYARDAFTDSALSAKVG EHKKDLPEWDAGSLTASELELNDVSGNDPNDMPRYNEENGVVSTYSSLSSTYTVLPRDNSBEFLKREARANQLABEIESAQYKARVALENDDRSEEEKYTAVORNSDR

gene

CDS

QY 596 SerHisProSerAlaHisGlySerProAlaProValSerThrMetProIysArgMetSer 615
DB 1722 TCTCAACCCCTCTGCTCATGGTTCTCAGCTCCTGTCTCTACTATGCCATAAAGCGATGCT 1781
QY 616 SerGluGlyProProArgMetSerProIysAlaGlnArgHisProArgAsnHisArgVal 635
DB 1782 TCAGAAGGACCCCCAAGGATGCTCCAAAGGCACAGCGCCACCCTCGGAATCACAGATC 1841
QY 636 SerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHisAsnProProSer 655
DB 1842 TCTGCTGGGAGAGCTCCATGCTAGTGGCTAGAAATTTGATCCCAAAATCCCCCAAGT 1901
QY 656 GluAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThrTrpSerSerVal 675
DB 1902 GAAGCAGCTGCTCCTCCAGTGGCAAGGACCAAGTCTCGAGGGGGAAACGTGGTCTCTCAGTG 1961
QY 676 ValSerGlyValProArgLeuSerProIysThrHisArgProArgSerProArgGlnAsn 695
DB 1962 GTCAGTGGGGTTCCAAAGGTATCTCCCAAACTCACAGCCCAAGTCTCCAGGCAGAGC 2021
QY 696 SerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAlaGlyIlePro 715
DB 2022 AGCATTTGGAACCTCTCCAGCGGGCTGTGCTTGTCTTCTCCCAAGCTGGCATCATCCCT 2081
QY 716 ThrGluAlaValAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSer 735
DB 2082 GCAGAAGCGGTTTCCATGCTCTCCCGCGCATCTCCGATCTCTGCTCCAGCCCTGCATCC 2141
QY 736 AsnArgAlaValThrProSerSerGluAlaIysAspSerArgLeuGlnAspGlnArgGln 755
DB 2142 AACAGAGCACTACCCCATCTATTGAGGCACAAAGATTCCAGCTTCAAGATCAGAGGCAG 2201
QY 756 AsnSerProAlaGlyAsnLysGluAsnIleLysProAsnGluThrSerProSerPheSer 775
DB 2202 AACTCTCTGTCAGGAGTAAAGAAATGTTAAAGCAAGTGAACATCATCCTAGCTTTTCA 2261
QY 776 LysAlaGluAsnLysGlyIleSerProValValSerGluHisArgLysGlnIleAspAsp 795
DB 2262 AAAAGCTGACAAAGGTATGTACACAGTTGTTTCTGAACACAGAAAAACAGATTGATGAC 2321
QY 796 LeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAsp 815
DB 2322 TTAAGAGAGTTTAAAGATGATTTTAGGTACAGCCCAAGCTCTACATCTGAATCTATGGAT 2381
QY 816 GlnLeuLeuAsnLysAsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIle 835
DB 2382 CNACTACTAAGCAAAAATAGAGAGGAGAAAAGTCCAGAGATTGATTAAGATAAACG 2441
QY 836 GluProSerAlaLysAspSerPheIleGlu-----AsnSerSerSerAsnCysThrSer 853
DB 2442 GAAGCAAGTGTAAAGATAGTTTTCATTGACAGCAGCAGCAGCAGCAACTGTATCCAGT 2501
QY 854 GlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnThrGluHis 873
DB 2502 GGCACGACGACCAACACAGCCCTAGCATCTCCCTTCATGCTTAGTAATGCAGAGCAC 2561
QY 874 LysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProAlaCysLysGln 893
DB 2562 AAGAGGGGCTGAGGTACATCCCAAGGGTGCAGACTTCCAGCCCAAGCTTGCANAA 2621
QY 894 GluLysAspAspLysGluLysLysAspAlaAlaGluGlnValArgLysSerThrLeu 913
DB 2622 GAGAAAGGATGACAGAGAAGAGAAGAACACAAACAGAGCAGGTTAGGAAATCGACATTG 2681
QY 914 AsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProLysProSerThrThr 933
DB 2682 AATCCCAATGCAAGAGGATTCACCCCTCGTCTTTCTCTCAGCCAAAGCCTTCTACTACC 2741
QY 934 ProThrSerProArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnPro 953
DB 2742 CCNAGGTCACTCGGCTTCNAGCACAAACCCAGCCCATCTATTGGTGGGTTCATCAGAGCCA 2801

QY 954 ThrProValTyrThrGlnProValCysPheAlaProAsnMetMetTyrProValProVal 973
DB 2802 GCTCCAGGTACACTCAGCCTGTGTGCTTCGCACCAATATGATGTATCCCGTCCCAAGT 2861
QY 974 SerProGlyValGlnProLeuTyrProIleProMetThrProMetProValAsnGlnAla 993
DB 2862 AGCCCGGGCGTACAAACCTTTATACCCCAATACCTATGACGCCCATGCTGTGAACAAGCC 2921
QY 994 LysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGlnAspGlnHisHis 1011
DB 2922 AAGACATATAGAGCAGGTAAAGTACCAAAATATGCCCCCAACAGCAGCAACACATCAT 2981
QY 1012 GlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThrProPro 1031
DB 2982 CAAAGCACCATGATGACCCAGCCTCCGCGCAGGGCCACCATCGTAGCCACCACCGGCC 3041
QY 1032 AlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsnGlnProLeuVal 1051
DB 3042 GCTTACTCCACTCAGTACGTTGCTCAGGCCCTCAGCAGTTTCCCAATCAGCCTTTGGTC 3101
QY 1052 GlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGly 1071
DB 3102 CAGCATGTGCGCATTATCAGTCTCAGCATCCTCATGTGTACAGTCTGTCTATACAAGT 3161
QY 1072 AsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerAla 1091
DB 3162 AATGCCAGGATGATGGACCAACCAGCAGCATGCTCAGCCTGGTGTGTCTTCTTCAAGT 3221
QY 1092 ThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyr 1111
DB 3222 GCTCAGTTCGGGGCTCAGAGCAGCAGCAGCAGCATGTATGTCATGTCCTCCAAATACCATAC 3281
QY 1112 AsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGln 1131
DB 3282 AACAAAGGAGACAAAGCCCTCTTTCTACTTTGGCCATTTCCACCGGCTCCCTCGCTCAGCAG 3341
QY 1132 TyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThr 1151
DB 3342 TATGCATATCTTAATGCGGCTCGCATCCACATATCCCCATCTCAGCCTTCGGCCACT 3401
QY 1152 ProThrGlyGlnGlnGlnSerGlnHisGlySerHisProAlaProSerProValGln 1171
DB 3402 CCCACCGACAGCAGCAAGAACCCAGCAGTGGTGAAGTCACTCCCTGCACCCAGCTGCTTTCAG 3461
QY 1172 HisHisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerProGlnGlnSerAla 1191
DB 3462 CACCATCAGCACCAAGGCTGCCAGGCTCTTCATCTGGCCAGTCCACAGCAGCAGTCCGCC 3521
QY 1192 IleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProAlaSerAsnThrGln 1211
DB 3522 ATTTATCATGCGGGCTGGCACCAACACCACCTTCCATGACACCTGCTCTTAATACACAG 3581
QY 1212 SerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIleHisProSerHis 1231
DB 3582 TCTCCACAGACAGCTTTCCAGCAGCAGCAACAGACAGTCTTTCACCATCCACCTTCTCAT 3641
QY 1232 ValGlnProAlaTyrThrAsnProHisMetAlaHisValProGlnAlaHisValGln 1251
DB 3642 GTTCAGCGGGATACACCAACCCCAACCCACATGCCCCCAGTACCTCAGGCTCATGTACAG 3701
QY 1252 SerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeuMetThrThrGln 1271
DB 3702 TCAGGAATGGTTCCTTCTCATCCAACTGCCCATCGCCATGATGCTAATGACGACACAG 3761
QY 1272 ProProGlyGlyProGlnAlaLeuAlaGlnSerAlaLeuGlnProIleProValSer 1291
DB 3762 CCACCC---GGTCCCAAGCGCCCTCGCTCAAAAGTAGTGACCTACAGGCCATTTCCAGTTTCG 3818
QY 1292 ThrThrAlaHisPheProTyrMetThrHisProSerValGlnAlaHisHisGlnGlnGln 1311
DB 3819 ACAACAGGCATTTCCCTTATATGACGACCCCTTCAGTACAAAGCCCAACCCACACAGCAG 3878
QY 1312 Leu 1312

```
Db      3879 TTG 3881
||||
RESULT 11
AR632099
LOCUS   AR632099      3798 bp      mRNA      linear      PAT 14-FEB-2005
DEFINITION Sequence 4 from patent US 6844431.
ACCESSION AR632099
VERSION   AR632099.1 GI:59773826
KEYWORDS
SOURCE    Unknown.
ORGANISM  Unknown.
REFERENCE Unclassified.
1 (bases 1 to 3798)
Pulat,S.M.
Nucleic acid encoding spinocerebellar ataxia-2 and products related
thereto
JOURNAL  Patent: US 6844431-A 4 18-JAN-2005;
Cedars-Sinai Medical Center; Los Angeles, CA
FEATURES
source   Location/Qualifiers
1..3798
/mol_type="mRNA"
ORIGIN
Alignment Scores:
Pred. No.: 5.04e-77 Length: 3798
Score: 5606.50 Matches: 1063
Percent Similarity: 94.20% Conservative: 41
Best Local Similarity: 90.70% Mismatches: 37
Query Match: 80.54% Indels: 31
DB: 6 Gaps: 8

US-10-802-228-2 (1-1312) x AR632099 (1-3798)
QY 148 ProArgProAlaAlaCysGluProValTyrGlyProLeuThrMetSerLeuLysProGlnGln 167
|||||
Db 11 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 70
|||||
QY 168 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 187
|||||
Db 71 CAG----- 73
QY 188 ProProProAlaAlaAlaAlaValArgLysProGlyGlySerGlyLeuLeuAlaSerPro 207
|||||
Db 74 ---CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 127
|||||
QY 208 AlaAlaAlaProSerProSerSerSerSerSerSerSerSerSerSerSerSerSerSer 227
|||||
Db 128 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 184
|||||
QY 228 SerValValAlaAlaAlaThrSer-----GlyGlyGlyArgProGlyLeuGlyArgGly 244
|||||
Db 185 GCGCGCGTGGCGTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 244
|||||
QY 245 ArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGlyIleTyrAlaAsn 264
|||||
Db 245 CGGAACAGTAGCAAGAGACTGCCTCAGCCTACGATTCTTTTGCATGGAATCTATGCAAAAC 304
|||||
QY 265 MetArgMetValHisIleLeuThrSerValValGlySerLysCysGluValGlnValLys 284
|||||
Db 305 GTGAGTAGTGTTCATATACCTACCTAGCTGTTGGATCGAAATGTGAAGTACAAAGTGAAA 364
|||||
QY 285 AsnGlyGlyIleTyrGluGlyValPheLeuThrTyrSerProLysCysAspLeuValLeu 304
|||||
Db 365 AACGAGGCGATATATGAGAGAGTGTTTTAAACATACAGTCTCTAAGTGTGACTTGTACTT 424
|||||
QY 305 AspAlaAlaHisGluLysSerThrGluSerSerSerSerSerSerSerSerSerSerSerSer 324
|||||
Db 425 GATGCTGCATCAGAAAAGTACAGAAATCCAGTTCGGGCGCAAAACGCGTGAAGAAATAATG 484
|||||
QY 325 GluSerIleLeuPheLysCysSerAspPheValValGlnGlnPheLysAspMetAspSer 344
|||||
```

485 GAGAGTGTGTTTGTTCAAATGCTCAGACTTCGTGTGGTGTACAGTCTTAAAGATACAGACTCC 544

345 SerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLysValAsnGlyGlu 364

545 AGTTATGACGAGAGAGATGCTTTTACTGACTCTGCTCTCAGCGCAAGGTGAATGGTGAG 604

365 HisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAlaAsnGluGluLeu 384

605 CACAAGGAGAAGGAGCTGGAGCCCTGGGATCGAGGGAGAGTCCACGCCAGCAGGAGCTG 664

385 GluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMetPheArgTyrAsn 404

665 GAG---CTGGAGATGATGTGTCTAATGATGGGACCCCAATGACATGTTTCGATATAAT 721

405 GluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSerTyrThrValPro 424

722 GAAGAGAATATGTTGGTGTGTCACATATGATAGCAGTTTATCTTCATATACGGTTCCT 781

425 LeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAlaAsnGlnLeuAla 444

782 TTAGAAAGGACAACTCAGAAAGATTTCTTAAACGGGAGGCAAGGCCAACCCAGTTAGCA 841

445 GluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGluAsnAspArg 464

842 GAAGAAATTTGAATCCAGTGTCTCAGTACAAAGCTCGTGTCCCTTTGAGAATGATCACCGG 901

465 SerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArgGluGlyHisSer 484

902 AGTGAGGAAGAAAAATACACAGCAGTCCAGAGAAACTGCAGTGACCGGGAGGGCATGGC 961

485 IleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArgGluValIleSer 504

962 CCCAACACTAGGACAAATAATATTTCTCTGGACAAAGAACAGAGAACTCTATCC 1021

505 TrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySerGlySerMetPro 524

1022 TGGGGAAGTGGGAGACAGAGCTCACACAGATGGGCCAGCCTGGGCCAGGCTCCATGCCG 1081

525 SerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySerAspGlnArgVal 544

1082 TCAAGAGCTGCTCTCACACTTTCAGATTTCACACCCGACGCTGGCTCAGACCAAGAGTA 1141

545 ValAsnGlyValProTyrProSerProCysProSerProSerSerArgProProSer 564

1142 GTTAATGGAGGTGTTCCCTGGCCATCGCTTGCCTCATCTCATCTCTCGCCACCTTCT 1201

565 ArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrProThrArgProPro 584

1202 CGCTACCACTAGTCCCAACTCTCTTCCACCTCGGCGACGCCACCATACACGGCGCGCC 1261

585 SerArgProProSerArgProSerArgProProSerHisProSerAlaHisGlySerPro 604

1262 TCCAGGCCCGCCCTCGAGGCGCATCCAGACCCCGCTCTCAGCCCTCTGCTCATGTGTTCCA 1321

605 AlaProValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSerPro 624

1322 GCTCTCTCTACTATGCTTAAACGATGTCTTTCAAGAGGACCCCAAGGATGTCTCCA 1381

625 LysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSer 644

1382 AAGGCACAGGCCACCCCTCGGAATCACAGATCTCTGCTGGGAGAGGCTCCATGTCTAGT 1441

645 GlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrProProValAlaArg 664

1442 GGCTAGAAATTTGTATCCCAATCCCAAGTGAAGCAGCTGCTCTCTCAGTGGCAAGG 1501

665 ThrSerProSerGlyThrTrpSerSerValValSerGlyValProProArgLeuSerPro 684

1502 ACCAGTCTTCAGGGGGAACGTGGTCTCTCAGTGGTGGGTTTCAAGGTTATCTCCC 1561

685 LysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyPro 704

1562 AAAACTCACAGACCCAGGCTCTCCAGGCGAGGAGCATTTGGAACATCTCCAGCGGCGCT 1621


```
QY 705 ValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAlaMetProIlePro 724
DB 1622 GTGCTTGCTTCTCCCAAGCTGGCATCATCCCTGCAGAAAGCGTGTTCATGCTGTCCTCC 1681
QY 725 AlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGlu 744
DB 1682 GCCGATCTCCCACTCTCCGACGCCCTGCATCCAAAGAGCACTGACCCCATCTATTAGG 1741
QY 745 AlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsn 764
DB 1742 GCAAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGTAGGGAGTAAGAAAT 1801
QY 765 IleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerPro 784
DB 1802 GTTAAAGCAAGTGAACATACCTAGCTTTTCAAAAGCTGACAAACAAGGTATGTACCA 1861
QY 785 ValValSerGluHisArgLysGlnIleAspAspLeuLysPheLysAsnAspPheArg 804
DB 1862 GTTGTTCTGAACACAGAAACAGATTGATGACTTAAAGAAAGTTTAAAGATGATTTTAGG 1921
QY 805 LeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGly 824
DB 1922 TTACAGCCAAGCTCTACATCTGAATCTATGGATCAACTACTAAGCAAAATAGAGAAGA 1981
QY 825 GluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysAspSerPheIle 844
DB 1982 GAAAAGTCACGAGATTTGATTAAAGATAAAACGGAAGCAAGTGTCTAAGGATAGTTTCAT 2041
QY 845 Glu-----AsnSerSerAsnCyThrSerGlySerSerLysProAsnSerProSer 862
DB 2042 GACACAGCAGCAGCAGCAGCAACTGTACCTAGTGGCAGCAGCAAGCAACAGCCCTAGC 2101
QY 863 IleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThrSerGln 882
DB 2102 ATCTCCCTTCCATCTGTTAGTATCGACGACACAGAGGGGGCTGAGGTACATCCCAA 2161
QY 883 GlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluLysLys 902
DB 2162 GGGGTGCAGACTTCCAGCCAGCTTCGCAAAACAAGAGAAGGATGACAGAGAAGAAGAAA 2221
QY 903 AspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPheAsnPro 922
DB 2222 GACACAAACAGACAGCGGTAGGAAATTCGACATTTGAATCCCAATGCAAGAGGATTCACCC 2281
QY 923 ArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGlnAlaGln 942
DB 2282 CGTTCTTCTCAGCCCAAGCGCTTCTACTACCCCAACGTCACCTCGGCTCAGACCAA 2341
QY 943 ProSerProSerMetValGlyHisGlnGlnProThrProValTyThrGlnProValCys 962
DB 2342 CCCAGCCCATCTATGGTGGGTCTCATCAGCAGCCAGCTCCAGTGTACACTCAGCGTGTGTC 2401
QY 963 PheAlaProAsnMetMetTyThrProValProValSerProGlyValGlnProLeuTyThrPro 982
DB 2402 TTCGCACCCCAATATGATGATCCCGTCCAGTAGCCGCGGCGTACACCTTTATACCACA 2461
QY 983 IleProMetThrProMetProValAsnGlnAlaLysThrTyArgAla-----ValPro 1000
DB 2462 ATACCTATGACGCCCATCGCTGTGAACCAAGCCACAGACATATAGAGCAGGTAAGATACA 2521
QY 1001 AsnMetProGlnGlnArgLysAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
DB 2522 AATATGCCCAACAGCAGCAAGACCACATCATCAAAGCACCATGATGATGATGATGATGAT 2581
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyThrSerThrGlnTyThrValAlaTy 1040
DB 2582 GCGGAGGGGCCACCCATCGTAGCCACCCCGCGCTTACTCCACTCAGTACGTGTCCTAC 2641
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyThrGlnSerGln 1060
DB 2642 AGCCCTCAGCAGTGTCCCAATCAGCCTTTGGTCCAGCATGTGCCCATATATCAGTCTCAG 2701
```

```
QY 1061 HisProHisValTySerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
DB 2702 CATCTCATGTGTAGTACAGTCTGTTCATACAAGGTAATGCCAGGATGATGGCACCCACGCA 2761
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyThrGlyAlaHisGluGlnThr 1100
DB 2762 CATGCTCAGCCTGGTGTAGTGTCTTTCAGCTGCTCAGTTCGGGGCTCAGAGCAGACG 2821
QY 1101 HisAlaMetTyAlaCysProLysLeuProTyThrAsnLysGluThrSerProSerPheTy 1120
DB 2822 CACGCCATGTATGCAATGTCCTCCAAATTACCATAACAAGGAGACAAGCCCTTCITCTTAC 2881
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyThrAlaHisProAsnAlaThrLeuHis 1140
DB 2882 TTTGCCATTTTCACCGGCTCCTCGCTCAGCAGTATGCACATCCTAATATGCGCGCTGCA 2941
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
DB 2942 CCACATACTCCCACTCCTCAGCCTTCGGCCACTCCACCGGACAGCAGCAAGCCAGCAT 3001
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
DB 3002 GGTGGAAGTCACCTCGACCCAGTCTGTTCAGCACCATCATCAGCAGGCTGCCAGGCT 3061
QY 1181 LeuHisIleAlaSerProGlnGlnSerAlaIleTyHisAlaGlyLeuAlaProThr 1200
DB 3062 CTTTCATCTGGCAGTCCACAGCAGCAGTGGGCCATTTATCATGCGGGCTGGCACCAACA 3121
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
DB 3122 CCACCTTCCATGACACCTGCTCTAATACACAGTCTCCACAGACGAGTTTCCCAGCAGCA 3181
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyThrAsnProPro 1240
DB 3182 CAACAGACAGTCTTCACCATCCACCTTCTCATGTTTCAGCCGGGATACACACCCACCC 3241
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3242 CACATGGCCCAAGTACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCCAAC 3301
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
DB 3302 GCCCATGGCCCAATGATGCTAATGACACACAGCCACCC-----GGTCCCAAGGCCGCCCTC 3358
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyMetThr 1300
DB 3359 GCTCAAGTGCACCTACAGCCCATTCAGTTCGACACAGCGCATTTCCCTTATATGAGC 3418
QY 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312
DB 3419 CACCTTTCAGTACAAAGCCCAACCAACAGCAGTTG 3454
```

RESULT 12

```
LOCUS BC097692 2558 bp mRNA linear VRT 02-JUN 2005
DEFINITION Xenopus laevis cDNA clone MGC:115230 IMAGE:5079818, complete cds.
ACCESSION BC097692
VERSION BC097692.1 GI:66910767
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE
AUTHORS 1 (bases 1 to 2558)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
Dev. Dyn. 225 (4), 384-391 (2002)
JOURNAL
PUBMED 12454917
REFERENCE
AUTHORS 2 (bases 1 to 2558)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
```


QY	449	SerSerAlaGlnTyrLysAlaArgValAlaLeuGluAenAspArgSerGluGlu	468
DB	957	TCAAGTCCCAAGTATAAAGCTAGAGTTGCTTTGGAAATGATGAGCGATCAGAGAGGAA	1016
QY	469	LysTyrThrAlaValGlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArg	488
DB	1017	AAATACACGGCTGTTCACAGGCCCTTAATGATCCAGGAAGGCACCTCACTGAATACCCAGA	1076
QY	489	GluAsnLysTyrIleProProGlyGlnArgAsnArgGluValIleSerTrp-----Gly	506
DB	1077	GAATATAGTACATTCACCTGTCGAGAGAAATAAGATGCAATGTTATGGACTACTCAC	1136
QY	507	SerGlyArgGlnAsnSerProArgMetGlyGlnProGlySerGlySerMetProSerArg	526
DB	1137	AGTGAAGGCAGAAATCTCTGCAAGGATAGACAGCTGCGATCAGGACCCCATTCCTTCAAGG	1196
QY	527	SerThrSerHisThrSerAspPheAsnProAsnSerGlySerAspGlnArgValValAsn	546
DB	1197	ACAGGATCACATTCCTTGACTACAAACCCCAATGCTGGAGCTGATCAGAGAGTTGTAAAT	1256
QY	547	GlyGlyValProTrpProSerProCysProSerProSerSerArgProProSerArgTyr	566
DB	1257	GGAGTGTTCCCTGGCCATCGCTTGTCAATCTCTTCCCGCCCATCTTCCTCGCTAC	1316
QY	567	GlnSerGlyProAsnSerLeuProProArgAlaAlaThrProThrArgProProSerArg	586
DB	1317	CAGTCAGGTCCCAACTCTCTCCACTCGGGCAGCACCCCTTACAAGGCGCCCTCAAGG	1376
QY	587	ProProSerArgProSerArgProProSerHisProSerAlaHisGlySerProAlaPro	606
DB	1377	CCCCCTTCGGCGCCATCAGGCCCCCGTCTCACCCCTCTGCTCATGGTTCCTCCAGCTCT	1436
QY	607	ValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSerProLysAla	626
DB	1437	GTCTCTACTATGCTTAAGCATGTCTTCAGNAGCACTCAAGATGTCTCCTTAAGCA	1496
QY	627	GlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSerGlyLeu	646
DB	1497	CAACGTCTACTCTCGTACTCACAGAGTTTCAATGCGAGGGGCAACATGTGTGGAGATTA	1556
QY	647	GluPheValSerHisAsnProProSerGluAlaAlaThrProProValAlaArgThrSer	666
DB	1557	GAATTTGTTCTCGCAATGAAGCTGGAGAGCGTCTGTGCAACCTGTGGCCGGAATAGT	1616
QY	667	ProSerGlyGlyThrTrpSerSerValValSerGlyValProArgLeuSerProLysThr	686
DB	1617	TCCTCAGGTGGACATGCTCTCTGAGCGGAGTTCAAGGTTATCCCAAGACA	1676
QY	687	HisArgProArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyProValLeu	706
DB	1677	CACAGGCCAAGTCTCCGCG-----CTAGGAAGTCTGACACCAAGTTCAAATGCG	1727
QY	707	AlaSerProGlnAlaGlyIleIleProThrGluAlaValAlaMetProIleProAlaAla	726
DB	1728	ACTTCGCCCCAGCTGGGACAGTGCCTCGGAGTCTGTTTATGCGAGTCCCTGCCCA	1787
QY	727	SerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGluAlaLys	746
DB	1788	TCCCTTACACCTGTAGCCCTGCATCTGATAGGCTGTACTCCCTCTGTGGACAAA	1847
QY	747	AspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnLys	766
DB	1848	GATACAAGACTACAAGATCAGACAGAAAATCCCCAGCTGGTGTGTAAGAAAATAGTAAA	1907
QY	767	ProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerProValVal	786
DB	1908	TACAGCGAATCT-----AGCAAACTGAAACAAAGCAAGTATCTCCAATGCT	1955
QY	787	SerGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsnAspPheArgLeuGln	806
DB	1956	TCTGAACAAAGAAACAGCTTGATGATTTAAAGAAATTTAAAGATGATTTTAGGCTACAG	2015
QY	807	ProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGly---Glu	825
DB	2016	CCCAGCTCAAGTCTCTGATGCTTTGGATCAGTTACCCGTCAAAAGTCGGGATAGTGACAA	2075
QY	826	LysSerArgAspLeuIleLys-AspLysIleGluProSerAlaLysAspSerPheIleGlu	845
DB	2076	AAGCAAGGGATCCGTTGAAGAAAAAATGGATTCAAATGCTAAAGAAATCAACAAGTGA	2135
QY	845	uAsnSerSerSerAsnCysThrSer---GlySerSerLysProAsnSerProSerIleSe	864
DB	2136	AAGCAGCAGTAATGCCCACTAGCAACAACAGCAGCAAGCCATGCAGTCCAGCATTTTC	2195
QY	864	rProSerIleLeu---SerAsnThrGluHisLysArgGlyProGluValThrSerGlnGlu	883
DB	2196	ACCATCTATTATCAGCAGCAGTCTCTGAACAACAAGAGAGGTCAGAGGTAACATCACAAG	2255
QY	883	yValGlnThrSerSerProAlaCysLysGlnLysAspAspLysGluGlu---LysLys	902
DB	2256	TGTACAGACCTCTGGGCTTC---AAACAGGATCGTATGACAAGGATGACATGAAGGA	2312
QY	902	sAspAlaAlaGlu 906	
DB	2313	AAATGCTGCAGAG 2325	
RESULT 13			
AK095017			
LOCUS	2766 bp	mRNA	linear
DEFINITION	AK095017	2766 bp	linear
ACCESSION	AK095017.1	GI:21754198	
VERSION	AK095017.1	GI:21754198	
KEYWORDS	oligo capping; fis (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
	Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,		
	Wakamatsu, A., Hayaishi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,		
	Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,		
	Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,		
	Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,		
	Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,		
	Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,		
	Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,		
	Tanikawa, M., Yanazaki, M., Ninomiya, K., Ishibashi, T., Yanashita, H.,		
	Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,		
	Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,		
	Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,		
	Hara, H., Tanase, T., Nomura, Y., Toqiya, S., Komai, F., Hara, R.,		
	Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,		
	Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,		
	Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,		
	Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,		
	Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,		
	Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,		
	Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,		
	Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,		
	Ozaki, K., Hiro, M., Omori, Y., Kawabata, A., Hiki, T., Kobatake, N.,		
	Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,		
	Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunura, K.,		
	Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,		
	Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,		
	Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,		
	Negase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,		
	Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.		
	Complete sequencing and characterization of 21,243 full-length		
	human cDNAs		
TITLE	Nat. Genet. 36 (1), 40-45 (2004)		
JOURN	14702039		
PUBMED	2		
REFERENCE	Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,		


```
Db 1236 TAGGAGTGGGAAGAAAATACACAGAGCTTCAGAGAAATTCAGTGACGTGAGGGGCA 1295
Qy sSerIleAenThrArgGluAsnLysTyfileProProGlyGlnArgAenArgGluValI1 503
Db 1296 CAGCATAAACACTAGGGAATAATATATATCTCTCGACAAAGAAATAGAGAAGTCAT 1355
Qy eSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySerGlySerMe 523
Db 1356 ATCTCTGGGGAAGTGGGAGACAGAAATTCACCGGTATGGCCAGCCCTGGATCGGCTCCAT 1415
Qy tProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySerAspGlnAr 543
Db 1416 GCCATCAGATCCATCTCTCACACTTCAGATTCAACCCGAATTCGTGTTACAGACCAAG 1475
Qy gValValAenGlyGlyValProTrpProSerProCysProSerProSerSerArgProPr 563
Db 1476 AGTAGTTAATGGAGGTGTTCCCTGGCCATCGCTTGCCCATCTCCTTCTCTCGCCACCC 1535
Qy oSerArgTyGlnSerGlyProAsnSerLeuProProArgAlaAlaThrProThrArgPr 583
Db 1536 TTCTCGCTACCAAGTCAGGTCCCAACTCTCTCCACCTCGGCGAGCCACCCCTACACGGCC 1595
Qy oProSerArgProProSerArgProSerArgProProSerHisProSerAlaHisGlySe 603
Db 1596 GCCCTCGGGGCCCCCTCGCGGCCATCCAGACCCCGCTCTCACCCCTCTGCTCATGGTTC 1655
Qy rProAlaProValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSe 623
Db 1656 TCCAGCTCTGCTCTACTATGCCCTTAACGATGTCCTTCAGNAGGCCCTCCNAGGATGTC 1715
Qy rProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSe 643
Db 1716 CCCAAAGGCCAGCGACATCTCGAAATCACAGAGTTTCTGCTGGCGAGGGTTCATATC 1775
Qy rSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrProProValAl 663
Db 1776 CAGTGGCTAGAAATTTGATCCCAACCCACCCAGTGAAGCAGCTACTCTCCAGTAGC 1835
Qy aArgThrSerProSerGlyGlyThrTrpSerSerValValSerGly 678
Db 1836 AAGGACCACTCCCTCGGGGGNACCTGGTGCATCAGTGTGAGTGGG 1881

RESULT 14
AR447281
LOCUS AR447281 1257 bp mRNA linear PAT 20-FEB-2004
DEFINITION Sequence 4 from patent US 6673535.
ACCESSION AR447281
VERSION AR447281.1 GI:42675576
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 1257)
  Fulst,S.M.
  TITLE Methods of detecting spinocerebellar Ataxia-2 nucleic acids
  JOURNAL Patent: US 6673535-A 4 06-JAN-2004;
  Cedars-Sinai Medical Center; Los Angeles, CA
FEATURES
  source
    1..1257
    /organism="unknown"
    /mol_type="mRNA"

ORIGIN
Alignment Scores:
Pred. No.: 3 67e-21 Length: 1257
Score: 1876.00 Matches: 365
Percent Similarity: 89.77% Conservative: 21
Best Local Similarity: 84.88% Mismatches: 18
Query Match: 26.95% Indels: 26
DB: 6 Gaps: 5

us-10-802-228-2 (1-1312) x AR447281 (1-1257)
```

```
Qy GlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 176
Db GGGCCGCTCACCATGTGCTGAAGCCGAGCCGAG----- 43
Qy GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
Db -----CCGCCCGCCGCCGACTGGCGCC 67
Qy LysProGlyGlySerGlyLeuAlaSerProAlaAlaProSerProSerSerSer 216
Db AAGCCCGCGCC-----GGCCTGCTCTCGTCGCCCGCCGCCCGCCGCGC---TCGCCCGCCG 12;
Qy SerValSerSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSer 234
Db ----- 181
Qy ---GlyGlyGlyArgProGlyLeuGlyArgGlyArgGlyArgGlyLeuProGln 253
Db CGCGCGCGCGCGCGCTCCCGCGCTGGCGAGAGTCCGAACAGTAGCAAGGAGACTGCCTCAG 241
Qy SerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSer 273
Db CCTACGATTTCTTTTGATGGAATCTATGCAACGCTGAGGATGGTTCTATATCTTACGTCA 301
Qy ValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrgluGlyValPhe 293
Db GTTGTGGATCGAAATGTGAAGTACAAGTGAACGAGGAGCATATATGAAGGAGTTTTT 361
Qy LysThrTySerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGlu 313
Db AAAACATACATCAGTCTCCTAGTGTGACTTGTGTTGTTGTCACATGAGAAAAGTACAGAA 421
Qy SerSerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp 333
Db TCCAGTTCCGGGGCCAAACGTAAGAAATAATGAGAGAGTGTTTGTTCAAATGCTCAGAC 481
Qy PheValValValGlnPheLysAspMetAspSerSerTyraLysArgAspAlaPheThr 353
Db TTCGTTGTGTGTACAGTTTAAAGATACAGACTCCAGTTTATGTCACGAGAGATGCTTTTACT 541
Qy AspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrp 373
Db GACTCTGCTCTCAGCGCAAGGTGAATGCTGAGCACAAGGAGAGGAGGAGGAGGAGGAGG 601
Qy AspAlaGlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAsn 393
Db ----- 658
Qy GlyTrpAspProAsnAspMetPheArgTyraAsnGluGluAsnTyrglyValValSerThr 413
Db GGATGGGACCCCAATGACATGTTTCGATATAATCAAGAGAATATGTTGTTGTTGTTCCACA 718
Qy TyrAspSerSerLeuSerSerTyThrValProLeuGluArgAspAsnSerGluGluPhe 433
Db TATGATAGCAGTTTATCTTATATACGGTTCCTTTAGAAAGGAGCAACTCAGAAAGATTT 778
Qy LeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTy 453
Db CTTAAACGGGAGGCAAGGCAACAGTTAGCAAGAAATTTGAATCCAGTGTCTCAGTAC 838
Qy LysAlaArgValAlaLeuGluAsnAspArgSerGluGluGluLysTyThrAlaVal 473
Db AAAGCTCGTGTGCGCCCTTGAGAAATGATGACCGGAGTGAAGGAGGAGGAGGAGGAGGAG 898
Qy GlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrrile 493
Db CAGAGAAATCTCAGTGAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 958
Qy ProProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerPro 513
Db CCTCTCGACAAAGAAACAGAGAAAGTCTCTATCTCTGGGAAAGTGGGAGAGAGAGGAGTCA 1018
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 23, 2005, 00:44:12 ; Search time 1532 Seconds

(without alignments)
5707.622 Million cell updates/sec

Title: US-10-802-228-2

Perfect score: 6961

Sequence: 1 MRSAAAPRSPAVATESRRF.....TAHPYMTSPVQAHQOOL 1312

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPIO_spool_p/US10802228/runat_21122005_110224_9692/app_query.fasta_1.1479
-DB=N Geneseq -QMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10802228@cgn 1 1 1232 @runat_21122005_110224_9692 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 21.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004as.*
- 13: geneseqn2004bs.*
- 14: geneseqn2005as.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6961	100.0	4481	2	AAV06552 Human SCA
2	6961	100.0	4481	2	Aaz23428 Human SCA
3	6961	100.0	4481	10	ADD18753 Human dis
4	6961	100.0	4481	10	ADH69466 Human spi

5	6961	100.0	4481	10	ABV75291	Human SCA
6	6961	100.0	4481	14	ADZ49164	Insulin s
7	6950.5	99.8	4367	2	AAV30270	Gene caus
8	6377	91.6	4200	2	AAT78912	Spinocere
9	6025.5	86.6	4225	10	ABV75290	Mouse SCA
10	5606.5	80.5	3798	2	AAV06553	Mouse SCA
11	2463	35.4	2006	3	AAE21628	Human bre
12	1876	27.0	1257	10	ABV75289	Mouse SCA
13	1816.5	26.1	4674	2	AAV90421	Human ata
14	1816.5	26.1	4674	8	ABZ72015	Human ata
15	1816.5	26.1	4674	10	ADH29078	Human chr
16	1257	18.1	1369	13	ADR08336	Full leng
17	1042.5	15.0	1707	6	ABX34831	Human cdn
18	1012.5	14.5	623	2	AAV17229	SCA2 gene
19	857	12.3	516	2	AAV06551	SCA2 gene
20	826	11.9	1226	6	ABX84413	Human cdn
21	799	11.5	485	9	ACH41056	Human foe
22	787	11.3	459	10	ADH69479	Human SCA
23	735	10.6	3210	4	ABL08977	Drosophil
24	725	10.4	5375	4	ABL08976	Drosophil
25	643	9.2	355	2	AAV17224	SCA2 gene
26	598	8.6	93631	11	ACN43984_4	Continuation (5 of
27	587	8.4	13416	10	AAV55815	Micromono
28	587	8.4	60196	10	AAV55810	Micromono
29	568.5	8.2	6492	12	ADK16023	Streptomy
30	568	8.2	30690	3	AAA92301	S. avermi
31	568	8.2	30690	4	AAH79277	Streptomy
32	566	8.1	11238	10	AAV55817	Micromono
33	564	8.1	88400	13	ADU47667	S. cyaneo
34	556.5	8.0	138203	14	ADZ51725	FR-008 po
35	543.5	7.8	85915	13	ADL15447	Streptomy
36	541.5	7.8	67167	14	ADY80051	Nucleotid
37	541	7.8	29467	14	ADY80051	Contig #1
38	535.5	7.7	13987	2	AAT80415	Hybrid sr
39	535.5	7.7	44377	2	AAT80414	Platenoli
40	535.5	7.7	44377	2	AAT78508	Platenoli
41	529	7.6	104096	13	ADX56092	Streptomy
42	528	7.6	65140	4	AAV17184	Streptomy
43	528	7.6	125401	4	AAV17186	Streptomy
44	520.5	7.5	73995	11	ACN43986	Human gen
45	520.5	7.5	84428	12	ADM45913	Streptomy

ALIGNMENTS

RESULT 1

AAV06552

ID AAV06552 standard; cdna; 4481 BP.

XX AC AAV06552;

XX AC AAV06552;

DT 06-JUL-1998 (first entry)

XX DE Human SCA2 cdna including CAG repeat region.

XX KW SCA2 gene; spinocerebellar ataxia-2; ataxin-2; human; diagnosis;
olivo-ponto-cerebellar atrophy; ss; ds.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

FT CDS 164..4101

FT primer_bind /*tag= a

FT primer_bind complement(631..648)

FT /*tag= b

FT repeat_region /note= "primer SCA2-A binding site"

FT repeat_region 658..723

FT /*tag= f

FT repeat_unit /note= "CAG repeat region"

FT repeat_unit 658..660

FT /*tag= g

FT repeat_unit /note= "CAG repeat"

FT repeat_unit 661..663

```
FT FT /tag= h
FT FT /note= "CAG repeat"
FT FT 664..666
FT FT /tag= i
FT FT /note= "CAG repeat"
FT FT 667..669
FT FT /tag= j
FT FT /note= "CAG repeat"
FT FT 670..672
FT FT /tag= k
FT FT /note= "CAG repeat"
FT FT 673..675
FT FT /tag= l
FT FT /note= "CAG repeat"
FT FT 676..678
FT FT /tag= m
FT FT /note= "CAG repeat"
FT FT 679..681
FT FT /tag= n
FT FT /note= "CAG repeat"
FT FT 685..687
FT FT /tag= o
FT FT /note= "CAG repeat"
FT FT 688..690
FT FT /tag= p
FT FT /note= "CAG repeat"
FT FT 691..693
FT FT /tag= q
FT FT /note= "CAG repeat"
FT FT 694..696
FT FT /tag= r
FT FT /note= "CAG repeat"
FT FT 700..702
FT FT /tag= s
FT FT /note= "CAG repeat"
FT FT 703..705
FT FT /tag= t
FT FT /note= "CAG repeat"
FT FT 706..708
FT FT /tag= u
FT FT /note= "CAG repeat"
FT FT 709..711
FT FT /tag= v
FT FT /note= "CAG repeat"
FT FT 712..714
FT FT /tag= w
FT FT /note= "CAG repeat"
FT FT 715..717
FT FT /tag= x
FT FT /note= "CAG repeat"
FT FT 718..720
FT FT /tag= y
FT FT /note= "CAG repeat"
FT FT 721..723
FT FT /tag= z
FT FT /note= "CAG repeat"
FT FT 740..757
FT FT /tag= c
FT FT /note= "primer SCA2-B binding site"
FT FT 899..900
FT FT /tag= e
FT FT /note= "predicted splice site"
FT FT 1070..1091
FT FT /tag= d
FT FT /note= "primer SCA2-14B binding site"
```

```
PR 19-JUL-1996; 96US-0022207P.
PR 08-OCT-1996; 96US-00727084.
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
XX Pulst S;
XX
XX WPI; 1998-086523/08.
DR P-PSDB; AAW33807.
XX
XX Nucleic acids encoding human and mouse ataxin 2 - a product of the
FT spinocerebellar ataxia 2 gene, SCA2; useful in the diagnosis of ataxia
FT type 2.
XX
XX Claim 6; Page 52-58; 98pp; English.
XX
XX This cDNA sequence corresponds to a novel SCA2 gene encoding a human
CC spinocerebellar ataxin-2 (SCA2) polypeptide, designated ataxin-2 (see
CC AAW33807). A trisomy 21 foetal brain cDNA library and an adult human
CC frontal cortex cDNA library in lambda ZapII were screened with probes
CC obtained by PCR amplification of plasmid AAP65122B (see AAV06551). PCR
CC products were used to screen the human adult frontal cortex library, and
CC 5' clones were obtained by RT-PCR of placental mRNAs. Overlapping clones
CC was used to generate the composite 481 bp sequence. Ataxin type 2 can be
CC diagnosed by detecting a genomic or transcribed mRNA sequence in an
CC individual having an expanded CAG repeat at a location corresponding to
CC the CAG repeat region of the SCA2 gene. The presence of at least 13 CAG
CC repeats above the normal level (22, occasionally 23, repeats) is
CC indicative of SCA2. Primers (see AAT99640-41) amplifying at least this
CC region are used for diagnosis. Also claimed are kits for detecting
CC mutations at the SCA2 locus, antisense oligonucleotides, and transgenic
CC animals useful for studying the physiological roles of ataxin-2 and its
CC effect upon behaviour
XX
XX Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 5.27e-200 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-10-802-228-2 (1-1312) x AAV06552 (1-4481)
QY 1 MetArgSerAlaAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 163 ATCGCTCAGCGCGCGCAGCTCCTCGGAGTCCCGGGTGGCCACCGAGTCTCGCGGCTTC 222
QY 21 AlaAlaAlaArgTTPProGlyTTPArgSerLeuGlnArgProAlaAlaArgSerGlyArg 40
Db 223 GCCGAGCCAGGTGGCCCGGGTGGCTCGCTCCAGCGCGCGCGCGGAGGGGGCGG 282
QY 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 283 GCGGCGGTGGCGCGCGCGCGGACCGGTATCCCTCCGCGCGCCCTCCCGCGCGGCC 342
QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db 343 GGCCCCCTCCCTCCCGCAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 402
QY 81 GlyAenGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100
Db 403 GGCACACGGCGGGCGCGGTTCGGCCCGGCTCCCGCGGCGCTCTTGGTCTCGGCGGGCT 462
QY 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
Db 463 CCGCGCCCTTTCGTCTCGTCTCTTCTCCCTCCCTCCGCGAGCCCGGGCGCCCTCCGCGCGG 522
QY 121 ProThrArgAlaSerProLeuGlyAlaAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 523 CCAACCGCGCGCTCCCGCTCGGCGCGCTCGTCTCCCGCGCGCGCTCCGCGGCTCTCTCTTG 582
```


Db	2743		CCAGCATTTCCCTTCAATCTTAGTAACCGAGACAAGAGGGACCTGAGGTCACT	2802
Qy	881	SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu	900	
Db	2803	TCCCAAGGGGTTACAGCTTCAGCGCCAGCATGTAAACAGAGAAAGACGATAGGAAGAG	2862	
Qy	901	LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe	920	
Db	2863	AAGAAAGACGCAGCTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGGAGTTC	2922	
Qy	921	AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln	940	
Db	2923	AAACCCAGGTTCTCTCTCAGCCAAAGCCCTTCTACTACCCCAACTTCACCTCGCGCTCAA	2982	
Qy	941	AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro	960	
Db	2983	GCACAACTAGCCCATCTATGGTGGGTCATCAACAGCCAACTCAGTTTATATCTACGCT	3042	
Qy	961	ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu	980	
Db	3043	GTTTGTTTTGACCAATATGATGTATCCAGTCCAGTCGAGCCAGCGGTCAACCTTTA	3102	
Qy	981	TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro	1000	
Db	3103	TACCAATATCTATGACGCCCATGCCAGTGAATCAAGCAAGACATATAGCAGGTACCA	3162	
Qy	1001	AsnMetProGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer	1020	
Db	3163	AAATGCCCCAACACGGCAAGACCAACATCATCAGAGTCCATGATGCACCCAGCGTCA	3222	
Qy	1021	AlaAlaGlyProProlIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr	1040	
Db	3223	GCACGGGGCCACCGATTGCAGGCCACCCACCGCTTACTCCAGCAATATGTTGCTTAC	3282	
Qy	1041	SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln	1060	
Db	3283	AGTCCTCAGCAGTTCCCAATCAGCCCTCTGTTCCAGCATGTGCCACATTCAGTCTCAG	3342	
Qy	1061	HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr	1080	
Db	3343	CATCCTCATGTCTATAGTCTCTGTAATACGGTAATGCTAGATGATGGCACCACCAACA	3402	
Qy	1081	HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr	1100	
Db	3403	CACGCCAGCGCTGGTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCG	3462	
Qy	1101	HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr	1120	
Db	3463	CATCCGATGTATGCATGTCCCAATTAACCATACAACAAGGAGACAAGCCCTCTTCTTAC	3522	
Qy	1121	PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis	1140	
Db	3523	TTTGCCATTTCCACGGGCTCCCTTGTCTCAGCAGTATGCGACCCCTAACGCTACCCCTGCAC	3582	
Qy	1141	ProHisThrProHisProGlnProSerSerAlaThrProThrGlyGlnGlnSerGlnHis	1160	
Db	3583	CCACATCTCCACACCTTCAGCTTCAGCTTACCCCTGACAGCAGCAGCAAGGCCACAT	3642	
Qy	1161	GlyGlySerHisProAlaProSerProValGlnHisHisGlnAlaAlaGlnAla	1180	
Db	3643	GGTGGAAAGTCATCTGCACCGCAGTCTCTGTTCCAGCACCAGCGCGCCCGCT	3702	
Qy	1181	LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr	1200	
Db	3703	CTCCATCTGGCCAGTCCACAGCAGCAGTCAGCCATTTACACCGGGGGCTTGGCGCAACT	3762	
Qy	1201	ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla	1220	
Db	3763	CCACCTTCCATGACACCTTGCTCCAAACACGAGTCCGCCACAGAAATGTTTCCAGAGCA	3822	
Qy	1221	GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro	1240	

Db	3823	CAACAGACTGCTTTTACGATFCCATCCTTCTCACGTTTCAGCGGCGGTATATCAACACCCACC	3888
Qy	1241	HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr	1260
Db	3883	CACATGGCCACGATGCTCTAGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCA	3942
Qy	1261	AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu	1280
Db	3943	GCCCATGGCGCAATGATGCTTAATGACACACACGCCACCCGGCGGTCCCCAGGCGGCCCTC	4002
Qy	1281	AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr	1300
Db	4003	GCTCAAAGTGCACTACAGCCCATTCACGTCCTGCACAACACAGCGCATTTCCCTTATATGACG	4062
Qy	1301	HisProSerValGlnAlaHisHisGlnGlnGlnLeu	1312
Db	4063	CACCTTCAGTACAGCCCAACACACAGCAGTTG	4098
RESULT 2			
AAZ23428			
ID	AAZ23428 standard; DNA; 4481 BP.		
XX			
AC	AAZ23428;		
XX			
DT	19-JAN-2000 (first entry)		
XX			
DE	Human SCA2 DNA.		
XX			
KW	Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC;		
KW	huntingtin polypeptide; Machado-Joseph disease; SCA1; SCA2; SCA6;		
KW	atrophin-1; cell death; apoptosis; Huntington's disease; head trauma;		
KW	Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke;		
KW	dentatorubropallidolysian atrophy; cell proliferation; cell survival;		
KW	neoplastic; malignant; autoimmune; fibrotic; ss.		
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
PH	163..4101		
FT	/*tag= a		
FT	/*tag= a		
FT	/*product= "SCA2"		
XX			
PN	WO9945944-A1.		
XX			
PD	16-SEP-1999.		
XX			
PF	11-MAR-1999; 99WO-US005250.		
XX			
PR	12-MAR-1998; 98US-00041886.		
XX			
PA	(BURN-) BURNHAM INST.		
XX			
PI	Bredesen DE, Rabizadeh S;		
XX			
DR	WPI; 1999-561617/47.		
DR	P-PSDB; AAY33495.		
XX			
PT	New proapoptotic dependence peptides, used to develop products for		
PT	treating, e.g. Alzheimer's disease.		
XX			
PS	Disclosure; Page 130-135; 1999p; English.		
XX			
CC	This invention describes novel pure proapoptotic dependence peptides		
CC	which comprise a sequence of an active dependence domain selected from		
CC	dependence polypeptides consisting of p75NTR, androgen receptor, DCC,		
CC	huntingtin polypeptide, Machado-Joseph disease gene product, SCA1, SCA2,		
CC	SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable of		
CC	inducing cell death and can be used to develop products to mediate or		
CC	inhibit apoptosis. The methods can be used for reducing the severity of a		
CC	proapoptotic dependence domain mediated pathological conditions e.g.		
CC	Huntington's disease, Alzheimer's disease, Kennedy's disease,		
CC	Spinocerebellar ataxias, dentatorubropallidolysian atrophy, Machado-		
CC	Joseph disease, stroke or head trauma. They can also be used for reducing		

CC the severity of a pathological condition mediated by upregulated cell
 CC proliferation or cell survival e.g. neoplastic, malignant, autoimmune or
 CC fibrotic conditions. This sequence encodes the human SCA2 polypeptide
 CC described in the method of the invention

XX
 SQ Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,27e-200 Length: 4481
 Score: 6961.00 Matches: 1312
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-10-802-228-2 (1-1312) x AA223428 (1-4481)

QY	1	MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgPhe	20
DB	163	ATGCGCTCAGCGCCGAGCTCTCGAGTCCGCGGTCCGCGGTCCGCGCTTC	222
QY	21	AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSerGlyArg	40
DB	223	GCGCGAGCAGGTGGCCGCGTGGCGTCTCCAGCGCGCGCGGCGGCGG	282
QY	41	GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro	60
DB	283	GCGCGCGTGGCGCGCCGCGACGATATCCCTCCGCGCGCCCTCCCGCGCGCC	342
QY	61	GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn	80
DB	343	GCGCCCGCTCCCTCCGCGCAGACTCGCTCCCTCCGCGCTCAGACTGT	402
QY	81	GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyGlyPro	100
DB	403	GGCAACGGCGCGCGCGGTTCGCGCCGCTCCCGCGCGTCTTGGTCTCGCGCGCT	462
QY	101	ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProAlaAla	120
DB	463	CCCCGCGCTTCGTCGTCGCTCTCTCCCTCCGCGCGCGCGCGCTCCGCGCGCG	522
QY	121	ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu	140
DB	523	CCAACCGCGCGCTCCCGCTCCGCGCGCGTCCGCTCCCGCGCGCTCCGCTTC	582
QY	141	AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr	160
DB	583	GCGCGCGCGCTCCCGCTGTCCCGCGCGCGTCCGCGCGGTATGGCGCGCTCACC	642
QY	161	MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	180
DB	643	ATGTCGCTGAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	702
QY	181	GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGly	200
DB	703	CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	762
QY	201	SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerSerSerSer	220
DB	763	AGCGGCTCTTAGCGTCCCGCGCGCGCTTCGCGCGCTTCGCTCTCGTCTCTCG	822
QY	221	SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly	240
DB	823	TCCTCGCGCCAGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	882
QY	241	LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrThrThrThrThr	260
DB	883	CTGGCGCAGGTCCGAACAGTAAACAAAGAGCTGCTCAGTCTCAGATTTCTTTGATGA	942
QY	261	IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu	280
DB	943	ATCTATGCAATATGAGATGGTTTATATATCTTACATCATGTTTGGCTCCAAATGTGAA	1002

QY	281	ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys	300
DB	1003	GTACAAAGTGAAATATGAGGTATATATGAGAGGTTTTTAAACCTTTACAGTCCGAAGTGT	1062
QY	301	AspLeuValLeuAspAlaAlaHisGlyLysSerThrGluSerSerSerGlyProLysArg	320
DB	1063	GATTTGGTACTTGTGCGCACATGAGAAAGTACAGAAATCCAGTTCGGGCGCGAAACGT	1122
QY	321	GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys	340
DB	1123	GAAGAAATAATGGAGAGTATTTTGTTCAAATGTTTCAGACTTTTGTGTGTCAGATTTAA	1182
QY	341	AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys	360
DB	1183	GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAA	1242
QY	361	ValAsnGlyGluHisLysGlyLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla	380
DB	1243	GTGAATGGCGAACACAAAGAGAGACCTGGAGCCCTGGGATGCGAGTGAACCTCACAGCC	1302
QY	381	AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet	400
DB	1303	AATGAGAACTTGAGGCTTTGGAAATGACGTATCTAATGGATGGGATCCCAATGATATG	1362
QY	401	PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer	420
DB	1363	TTTCGATATTAATGAAGAAATATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1422
QY	421	TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla	440
DB	1423	TATACAGTCCCTTAGAAGAGATAACTCAGAAGAAATTTTAAACCGGAAGCAAGGCA	1482
QY	441	AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu	460
DB	1483	AACCACTAGCAGAGAAATTCAGTCAAGTCCAGTACAAAGCTCGAGTGGCCCTCGAA	1542
QY	461	AsnAspAspArgSerGluGluGlyTyrThrAlaValGlnArgAsnSerSerGluArg	480
DB	1543	AATGATGATGAGGTAGGAGAGAAATACACAGCAGTTCAGAGAAATTCAGTGAACGT	1602
QY	481	GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg	500
DB	1603	GAGGGCAGCAGATAAACACTAGGGAATATAATATATTCCTCTCGACAAAGAAATAGA	1662
QY	501	GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer	520
DB	1663	GAAATCATATCTCTGGGAAAGTGGGAGACAGAAATTCACCGGTATGGCCACCGCTGATCG	1722
QY	521	GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer	540
DB	1723	GGCTCCATGCCATCAAGATCCACTTCCACACTTCAGATTTCAACCCGAAATTCGGTTCA	1782
QY	541	AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer	560
DB	1783	GACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCTTGCCTATCTCTCTCTCTCT	1842
QY	561	ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro	580
DB	1843	CGCCCACTTCTCGTACCAGTCAGGTCCCAACTCTCTTCCACTCGGCGAGCCACCCCT	1902
QY	581	ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla	600
DB	1903	ACACGCGCGCTCCAGGCGCCCTCGGCGCCATCCAGACCCCGCTCTCAACCCCTCTGCT	1962
QY	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
DB	1963	CATGGTTCTCCACTCTCTGTCTACTATGCTTAAACGCAATGTTTCAGAGAGGCGCTCCA	2022
QY	621	ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly	640
DB	2023	AGGATGTCCCAAGAGCCAGCAGCATCTCTCGAAATACACAGAGTTTCTGCTGGAGGGGT	2082
QY	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro	660

|||||
2083 TCATATCCAGTGGCTAGAAATTTGATATCCCAACACCACCCAGTGAAGCAGCTACTCCT 2142
QY
661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680
Db
2143 CCAGTAGCAAGACACAGTCCCTCGGGGGAACGTGGTCATCAGTGGTCAAGTGGGTTC 2202
QY
681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db
2203 AGATTATCCCTTAAACCTCATAGACCAGGTCTCCAGACAGAACAGTATTGGAAATACC 2262
QY
701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
Db
2263 CCAGTGGGCCAGTCTTCTGCTTCTCCCAAGCTGGTATTATTTCCAACTGAAGCTGTGCC 2322
QY
721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
Db
2323 ATGCCTATTCCAGCTGCATCTCTAGCGCTGCTAGTCTGCTGCATCGAACAGAGCTGTACC 2382
QY
741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
Db
2383 CCTTCTAGTGGAGCTTAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGAGG 2442
QY
761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
Db
2443 AATAAGAAAAATTAAACCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAACACAA 2502
QY
781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
Db
2503 GGTATATCACAGTGTCTTCTGAAATAGAAACAGATTGATTTAAAGNAATTTAAG 2562
QY
801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db
2563 AATGATTTTAGGTTCAGCCAAAGTTCTACTTCTGAATCTATGCATCAACTCTAAACAA 2622
QY
821 AsnArgGluGlyGluLysSerArgAspIleLysAspLysIleGluProSerAlaLys 840
Db
2623 AATAGAGGGGAGAAAAATCAAGAGATTGTATCAAGAAACAAAATTTGAACCAAGTCTAAG 2682
QY
841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Db
2683 GATTTCTTTCATTTGAAATAGCAGCAGCACTGTACCAAGTGGCAGCAGCAAGCCGAATAGC 2742
QY
861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db
2743 CCCAGCATTTCCCTTCAATCTTAGTAAACGAGGACACAGAGGGGACCTCAGGTCACT 2802
QY
881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900
Db
2803 TCCCAGGGGTTTCAGACTTCCAGCCCCAGCATGTAAACAGAGAAAGACGATTAAGGAAGAG 2862
QY
901 LysLysAspAlaAlaGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db
2863 AAGAAAGACGAGCTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGAGGTTC 2922
QY
921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db
2923 AACCCACGTTCTTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAA 2982
QY
941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrrThrGlnPro 960
Db
2983 GCACAACTTAGGCCCATCTATGTGGGTTCATCAACAGCCCAACTCCAGTATTATCTCAGCCT 3042
QY
961 ValCysPheAlaProAsnMetMetTyrrProValProValSerProGlyValGlnProLeu 980
Db
3043 GTTTGTTTGGACCAAAATATGATGATATCCAGTCCCGAGGCCCGCGGTGCAACTTTA 3102
QY
981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrrArgAlaValPro 1000
Db
3103 TACCCAAATACCTATGACGCCCATGTCAGGTGAATCAAGCCCAAGACATATAGACAGTACCA 3162
QY
1001 AsnMetProGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
|||||

Db
3163 AATATGCCCCAACAGCGGCAAGACCAGCATCATCAGAGTGCCATGATGCACCCAGCGTCA 3222
QY
1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrrSerThrGlnTyrrValAlaTyrr 1040
Db
3223 GCAGCGGGCCACCGATTGTCAGGCACCCACACAGCTTACTCCAGCAATATGTTGCCCTAC 3282
QY
1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrrGlnSerGln 1060
Db
3283 AGTCTCTCAGCAGTTTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAG 3342
QY
1061 HisProHisValTyrrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db
3343 CATCTCATGTCTATAGTCTGTAAATACAGGTTAATGCTAGAAATGATGGCACCAACA 3402
QY
1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrrGlyAlaHisGluGlnThr 1100
Db
3403 CACGCCACCGCTGTGTAGTATCTTCTCAGCAACTCAGTACGGGGCTCATGAGCAGACG 3462
QY
1101 HisAlaMetTyrrAlaCysProLysLeuProTyrrAsnLysGluThrSerProSerPheTyrr 1120
Db
3463 CATCGCATGTATGATGTCTCCAAATTTACCATCAACAAGGAGACAAGCCCTTCTTTCTTAC 3522
QY
1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrrAlaHisProAsnAlaThrLeuHis 1140
Db
3523 TTTGCCATTTCCAGGGGTCTCTTGTCTCAGCAGATATGCGCACCTTAAGCTACCTCGTGCAC 3582
QY
1141 ProHisThrProHisProGlnProSerSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db
3583 CCACATATCTCCACACCCCTCAGCTTACCTACCCCTCAGCAGCAGCAAGCAACAT 3642
QY
1161 GlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAla 1180
Db
3643 GGTGGAGTCTATCTCTGCACCCAGTCTGTTCAGCACCATCAGCACCCAGCGCCGAGGCT 3702
QY
1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrrHisAlaGlyLeuAlaProThr 1200
Db
3703 CTCCATCTGGCCAGTTCACAGCAGCATCAGCCATTTTACCAAGCGGGCTTGGCCCAACT 3762
QY
1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db
3763 CCACCTCCATGACACCTGCTCCCAACACAGCAGTCGCCACAGAAATAGTTTCCAGCAGCA 3822
QY
1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrrThrAsnProPro 1240
Db
3823 CAACAGACTGTCTTTACGATCCATCTCTCAGCTTACCGCGGCTATACCAACCCACCC 3882
QY
1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db
3883 CACATGGCCACAGTACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCTCATCCAACT 3942
QY
1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db
3943 GCCCATGCGCCAATGATGCTTAATGACGACACAGCCACCCCGCGGTCCCGAGCGCCCTC 4002
QY
1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrrMetThr 1300
Db
4003 GCTCAAGTGCACACTACAGCCCATTTCCAGTCTCGAACACAGCGCATTTTCCCTATATGACG 4062
QY
1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db
4063 CACCTTCAGTACAGAGCCCAACACAGCAGGTTG 4098

RESULT 3

ADD18753
ID ADD18753 standard; DNA; 4481 BP.
XX
AC ADD18753;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human disease related protein DNA sequence SeqID184.
XX
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;

KW	antiarteriosclerotic; vulnerary; gene therapy;
KW	hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
KW	inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW	glucose transportation; catecholamine synthesis; iron transport;
KW	nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW	retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW	inflammatory condition; wound healing; gene; ds.
XX	
OS	Homo sapiens.
OS	
PN	WO2003018621-A2.
XX	
PD	06-MAR-2003.
XX	
XX	23-AUG-2002; 2002WO-GB003892.
XX	
XX	23-AUG-2001; 2001GB-00020558.
PR	05-OCT-2001; 2001GB-00024037.
PR	
XX	(OXFO-) OXFORD BIOMEDICA UK LTD.
XX	
PA	Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
PI	
PI	WPI; 2003-290046/28.
XX	P-PSDB; ADD18752.
DR	
DR	
XX	New substantially purified polypeptide, useful for diagnosing or treating
PT	a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT	injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT	wound healing.
XX	
PS	Claim 27; SEQ ID NO 184; 424pp; English.
XX	
CC	This invention relates to novel human genes and gene product which are
CC	implicated in certain disease states. Compounds which modulate the
CC	proteins of the invention may have cytostatic, antiinflammatory,
CC	ophthalmological, antiarteriosclerotic or vulnary activities. The
CC	sequences of the invention may be useful for gene therapy. The invention
CC	may be useful for diagnosing or treating a hypoxia-regulated condition,
CC	such as tumorigenesis, angiogenesis, apoptosis, inflammation,
CC	erythropoiesis, or the biological response to hypoxia conditions
CC	including processes such as glycolysis, gluconeogenesis, glucose
CC	transportation, catecholamine synthesis, iron transport or nitric oxide
CC	synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC	injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC	inflammatory conditions or wound healing. The present sequence is that of
CC	a disease related protein encoding DNA sequence of the invention.
XX	
SQ	Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	5, 27e-200 Length: 4481
Score:	6961.00 Matches: 1312
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	10 Gaps: 0
US-10-802-228-2 (1-1312) x ADD18753 (1-4481)	
Qy	1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db	163 ATGCGCTCAGCGCGCAGCTCTCGAGTCCCGCGTGCGCCACCGAGTCTCGCGCTTC 222
Qy	21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db	223 GCCGACGACGAGTGGCGCGGCTCGCTCCAGCGCGCGCGCGCGAGCGGCGG 282
Qy	41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db	283 GCGCGCGTGGCGCGCGCGCGCGGACGTAFCCTCCGCCGCCCTCCCGCGCGCGCCC 342
Qy	61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80

Db 1423 TATACAGTGCCTTAGAAGAGATAACTCAGACGAATTTTAAACGGGAAGCAAGGCCA 1482
Qy 441 AenGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db 1483 AACCAAGTTAGCAGAGAAATTTAGTCAAGTGCCAGTACAAAGCTCGAGTGGCCCTGGAA 1542
Qy 461 AenAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db 1543 AATGATGATAGGAGTGAGAGAGAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGT 1602
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
Db 1603 GAGGGGCACAGCATAAACATAGGAGAAATAAATATATTCCTCTGGACAAAGAAATAGA 1662
Qy 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1663 GAAGTCATATCTCGGGAAGTGGAGACAGAATTCACCGCGTATGGCCAGCCTGGATCG 1722
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACATTCAGATTTCAACCCGCAATTTCTGGTTCA 1782
Qy 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCTGGCCATCGCCTTGGCCATCTCCTTCCTCT 1842
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1843 CGCCACCTCTCTCGTACCAAGTCAGGTCCCAACTCTCTTCACCTCGGCGACCACTCT 1902
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db 1903 ACACGGCGCCCTCCAGGCCCTCCAGGCCCTCCAGGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1963 CATGTTCTCCAGCTCCTCTCTACTATGCTTAACGCATGTCTTCAGAAAGGGCTCCA 2022
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Db 2023 AGGATGTCCCCAAAGGCCAGCGACATCTCGAATATCACAGAGTTTCTGCTGGAGGGGT 2082
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
Db 2083 TCCATATCCAGTGGCTAGAAATTTGATATCCCAACCAACCCAGTGAAGCAGCTACTCCT 2142
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680
Db 2143 CCAGTAGCAAGACCAAGTCCCTCGGGGGAGCGTGGTTCATCAGTGGTCAGTGGGGTTCCA 2202
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db 2203 AGATTATCCCTTAAACTCATAGACCCAGGCTCCACAGACAGACAGTATTGGAAATACC 2262
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIlelleProThrGluAlaValAla 720
Db 2263 CCCAGTGGGCCAGTTCTTGCTTCTCCCAAGCTGTTATTATTCCAACTGAAGCTGTTGCC 2322
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
Db 2323 ATGCCATTTCAGCTGCGATCTCTACGCGTGTAGTCTGTGATCGAATCGAAGAGCTGTACC 2382
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
Db 2383 CTTCTAGTAGGCTAAAGATTCCAGGCTTCAGATCAGAGGCAGAACTCTCCTGCAGGG 2442
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
Db 2443 AATAAAGAAAAATATTAACCCCAATGAAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
Db 2503 GGTATATCACCAGTTGTTTCTGAAATAGAAACAGATTGATTTTAAAGAAATTTTAAG 2562

Qy 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db 2563 AATGATTTTAGGTTACAGCCAAAGTTCTACTTCTGAATCTATGGATCAACTACTTAACAAA 2622
Qy 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
Db 2623 AATAGAGAGGAGAGAAAATCAAGAGATTTGTATCAAAAGCAAAAATTGAACCAAGTGCTTAAG 2682
Qy 841 AspSerPheIleGluAsnSerSerAsnSerSerThrSerGlySerSerLysProAsnSer 860
Db 2683 GATTCTTTTCATTGAAATATAGCAGCAGCAACTGTACCACTGGCAGCAGCAACCCGAATAGC 2742
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db 2743 CCCAGCATTTCCCTTCAATACTTAGTAACACGAGCAGCAAGAGGGGACCTGAGTCACT 2802
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnLysAspAspLysGluGlu 900
Db 2803 TCCCAAGGGGTTTCAGACTTCCAGGCCAGCATGTAAACAAGAGAAAGACGATAAGGAAG 2862
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 2863 AAGAAAGACGCGCTGAGCAAGTTAGAAATCAACATTGAATCCCAATGCAAAAGAGTTT 2922
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2923 AACCCAGTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCCCTCAA 2982
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db 2983 GCACAACTTAGCCCATCTATGGTGGTTCATCAACAGCCCACTCCAGCTTTATACTCAGCCT 3042
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db 3043 GTTTGTTTTGCACCAATATGATGATCCAGTCCAGTGAGCCAGGGCGTCAACCTTTA 3102
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 3103 TACCCAATACCTATGACGCCCATGCCAGTGAATCAAGCCAAGACATATAGAGCAGTACCA 3162
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
Db 3163 AATATGCCCAACAGCGGCAAGACAGCATCATCAGAGTGGCATGTCACCCAGCGTCA 3222
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 3223 GCAGGGGCCCCAGCATTTGACGCCACCCACAGCTTACTCCAGCAATATGTTGCCCTAC 3282
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 3283 AGTCCTCAGCAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCACATTAATCAGTCTCAG 3342
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCCTCATGTCTATAGTCTCTGTAATACAGGGTAAATGCTAGAAATGATGGCACCAACA 3402
Qy 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Db 3403 CACGCCACGCTGGTTTAGTATCTTTCACAACTCAGTACGGGGCTCATGACGACAGC 3462
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3463 CATGCGATGATGATGTGTCCCAATTACCATACAAAGGAGACAAAGCCCTTCTTCTTAC 3522
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTGCCATTTCCAGGGCTCCCTTGCTCAGCAGTATGCGCACCCCTTAACGCTACCTGAC 3582
Qy 1141 ProHisProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db 3583 CCACATCTCCACACCTCTCAGCCTTACGCTTACCCCTTACGACAGCAGCAAGCCNACAT 3642

QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB 1003 GTACAAAGTGAATAATGAGGTATATATCAAGAGATTTTAAAACTTACAGTCCGAAGTGT 1062
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
DB 1063 GATTTGGTACTTGTATGCCGCACATCAGAAAGATCAGAAATCCAGTTCGGGGCCGAAACGT 1122
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
DB 1123 GAAGAAATAATGGAGAGATTTTGTTCAAATGTTCCAGACTTTGTGTGTCAGTTTAAA 1182
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB 1183 GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAA 1242
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTTPAspAlaGlyGluLeuThrAla 380
DB 1243 GTGAATGGCGAACACAAAGAGAGAGACCTGGAGCCCTGGGATGCGAGGTGAACTCACGCC 1302
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTTPAspProAsnAspMet 400
DB 1303 AATGAGAACTTGGAGCTTTGGAAATGACGTATCTAATGATGGATCCCAATGATATG 1362
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB 1363 TTTTCGATATAATGAAGAAAAATATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB 1423 TATACAGTGCCCTTAGAAGAGATAACTCAGAAAGAAATTTTAAAAACGGGAAGCAAGGCA 1482
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1483 AACCAATTAGCAGAGAATAATGAGTCAAGTGCCCAAGTACAAAGCTCGAGTGGCCCTGGAA 1542
QY 461 AsnAspArgArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB 1543 AATGATGATAGCAGTGGAGAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGT 1602
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
DB 1603 GAGGGGCACAGCATAAACACTAGGGAAAAATAAATATATTCCTCTCGSACAAAGAAATAGA 1662
QY 501 GluValIleSerTTPGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB 1663 GAAGTCATATCTGGGGAAGTGGGAGACAGAAATTCACCGCGTATGGGCGAGCTGGATCG 1722
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATTCGTGTTCA 1782
QY 541 AspGlnArgValValAsnGlyGlyValProTTPProSerProCysProSerProSerSer 560
DB 1783 GACCAAGAGATGTTAAATGGAGGTGTTCCCTGGCCATCGCTTGGCCCATCTCTCTCTCT 1842
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
DB 1843 CGCCACCTTCTCGTACCAGTCAGTCCCAACTCTCTTCACCTCGGGGAGCCACCCCT 1902
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
DB 1903 ACACGGCCGCCCTCCAGGCCCCCTCGGGGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
DB 1963 CATGGTCTCCAGCTCTCTCTACTATGCTTAAACGCATGTCTTTCAGAAAGGGCCCTCA 2022
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
DB 2023 AGGATGTCCCCAAAGGCCCGCGACATCTCGAANAATCACAGATTTCTGCTGGGAGGGT 2082

QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
DB 2083 TCCATATCCAGTGGCTAGAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCTCT 2142
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTTPSerSerValValSerGlyValPro 680
DB 2143 CCAGTAGCAAGGACAGTCCCTCGGGGGAACGTGGTTCATCAGTGGTTCAGTGGGTTCCA 2202
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
DB 2203 AGATTATCCCTTAAAACTCATAGACCCAGGTCTCCAGACAGAACAGTATTTGGAAATACC 2262
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
DB 2263 CCCAGTGGCGCAGTCTTGTCTTCCCAAGCTGTATTAATTCCTCAACTGAAGCTTTGCC 2322
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGGCTATTCCAGCTGTCATCTCCAGCTCTAGTCTGTGATCGAACAGAGCTGTACC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2383 CTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGCAGG 2442
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2443 AATAAAGAAAAATATAAACCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
DB 2503 GGTATATCACCAGTGTGTTCTGAAACATAGAAAAACAGATTGATGATTTAAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 AATGATTTTAGGTTACAGCCAGTTCTACTTCTGATCTATGGATCAACTACTATAACAAA 2622
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2623 AATAGAGAGGGAGAAAAAATCAAGAGATTTGATCAAAAGACAAAAATTTGAACCAAGTCTAAG 2682
QY 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
DB 2683 GATTCTTTTCATTTGAAAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAAATAGC 2742
QY 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
DB 2743 CCCAGCATTTCCCTTCAATACTTAGTAACAGGAGCACAAAGGGGACCTGAGGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900
DB 2803 TCCCAAGGGGTTCAAGCTTCCAGCCAGCATGTAAACAAGAGAAAGACGATAAGGAAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
DB 2863 AAGAAAGACGCGCTGAGCAAGTTAGGAAATCAACATTGAATTCCTCAATGCAAGAGGATTC 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
DB 2923 AACCCACGTTCTTCTCTCAGCCAAAGCCTTCTACTACCCCACTTCACTCTCGGCTCAA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
DB 2983 GCACAAACCTAGCCCATCTATGGTGGGTTCATCAACAGGCCCACTCCAGCTTTATACTCAGCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
DB 3043 GTTTGTTTTCACCAAAATATGATGTATCCAGTCCCGAGTGGCCCGGCGTGCACCTTTA 3102
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
DB 3103 TACCCNATACCTATGACGCCCATTCAGGAGTGAATCAAGCCAGACATATAGAGCAGTACCA 3162
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020

Db 3163 AATATGCCCAACAGCGGCAAGCAGCATCATCAGATGCCATGATGATGCCAGCGTCA 3222
QY 1021 AlaaAGlyProPheAlaAlaThrProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 3223 GCAGCGGGCCCAACGATTCAGCCACCCACAGCTTACTCCACGCAATATGTGCCTAC 3282
QY 1041 SerProGlnGlnPheProGlnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 3283 AGTCCTCAGCAGTTCACCAATCAGCCCTGTGTTCAGCATGTGCCACATTCAGTCTCAG 3342
QY 1061 HisProHisValTyrSerProValIleGlnGlnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCCTCATGTCTATAGTCTGTGAATACAGGGTAATGCTAGATGATGACCAACCA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGlnGlnThr 1100
Db 3403 CACGCCAGCCTGGTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGAG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3463 CATGCATGTATGATGTCCTCCAAATPACCATACCAAGCAGACAGCCCTTCTTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTCGCATTTCCAGCGGCTCCCTTGCTCAGCAGTATGCGCACCTAACGCTACCCCTGCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
Db 3583 CCACATCTCCACACCTCAGCCTTCTCAGTACCCCTGAGCAGCAGCAGCAGCAGCAGC 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAla 1180
Db 3643 GTTGAAGTCATCTCCACCCAGCTCTGTTCAGCACCATCAGCACCGCGCCAGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCCATTTACCCAGCGGGCTTGGCCAACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3763 CCACCTCTCATGACATCTGCTCCAAACAGCAGTCCGACAGATAGTTTCCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3823 CAACAGCTGTCTTTACGATCCATCTCTCAGCTTCAGCGGCGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3883 CACATGCCCCAGCTACCTCAGGCTCATGTACAGTCAGGAATGTTCTCTTCATCCAACT 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrGlnProGlyGlyProGlnAlaAlaLeu 1280
Db 3943 GCCCATGCGCAATGATGTAAAGACACACGCCCGCGGCTCCCGAGCGGCCCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
Db 4003 GCTCAAGTGCATACAGCCATTCAGTCTCAGAACAGCGCATTTCCCTATATGAGC 4062
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db 4063 CACCTTTCAGTACAGCCACCCACCAACAGCAGTTG 4098

RESULT 5

ABV75291

ID ABV75291 standard; DNA; 4481 BP.

XX AC

XX AC

XX AC

DT 07-MAR-2003 (first entry)

DE Human SCA2 polypeptide encoding DNA.

XX

KW SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity;
KW memory impairment; mutant; human; gene; ds.

XX Homo sapiens.

FH Location/Qualifiers
FT 163..4101
FT /*tag= a
FT /product= "SCA2"

PN WO200289567-A2.

PD 14-NOV-2002.

XX 07-MAY-2002; 2002WO-US014698.

XX 07-MAY-2001; 2001US-0289231P.

XX (CEDA-) CEDARS SINAI MEDICAL CENT.

XX Pulst SM;

XX WPI; 2003-103467/09.

XX P-PSDB; ABB82698.

XX New mutant mouse, useful as a disease model for identifying agents for
XX treating obesity or memory impairment, comprises a disrupted
XX spinocerebellar ataxia 2 gene.

XX Disclosure; Page 65-72; 76pp; English.

XX The invention relates to a mutant non-human mammal having a disrupted
XX spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a
XX disease model for identifying agents or compounds that may treat obesity
XX or memory impairment. The stem cell is used for generating the mutant
XX animal. The present sequence represents the nucleotide sequence (GenBank
XX Accession No. U70323) encoding the human SCA2 polypeptide

XX Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,27e-200 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-802-228-2 (1-1312) x ABV75291 (1-4481)

QY 1 MetArgSerAlaAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20

Db 163 ATGCGCTCAGCGCGCCAGCTCTCGAGTCCGCGTGGCCACCGAGTCTCCCGCTTC 222

QY 21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgArgSerGlyArg 40

Db 223 GCGCAGCAGGTGGCGCGGTGGCTCGCTCAGCGCGCGCGCGCGCGCGCGCGCG 282

QY 41 GlyGlyGlyValAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60

Db 283 GCGCGGTGGCGCGCGCGCGCGCGGTATCCCTCCCGCGCGCTCCCGCGCGCGCC 342

QY 61 GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80

Db 343 GCG 402

QY 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100

Db 403 GCGCAGCG 462

QY 101 ProArgProPheValValValLeuLeuLeuProLeuAlaSerProGlyAlaProAlaAla 120

Db 463 CCGCGCGCGCTTCGTCGTCGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 522

QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB 523 CCAACCGCGCTCCCGCTCGGCGCGGTGCGTCCCGCGCGGTTCGCGGTCTCCTTG 582
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB 593 GCGCGCGCGCTCCCGGTGCTCCCGCGCGGTGCGAGCGGTGTATGGGCGCTCACC 642
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 643 ATGTCGCTGAAGCCCGACGACGACGACGACGACGACGACGACGACGACGACGACG 702
QY 181 GlnGlnGlnGlnGlnGlnProProProAlaAlaAlaAsnValArgLysProGlyGly 200
DB 703 CAGCAGCAGCAGCAGCAGCAGCGCGCGCGCGCGCTGCGCAATGTCCGCAAGCCCGCGCGC 762
QY 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
DB 763 AGCGGCGCTTCTAGCGCTCGCGCGCGCGCGCTTCCGCGTCCGCTCTCGGTCTCCTCG 822
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
DB 823 TCCTCGGCGACGGCTCCCTCTCGTGTGCGGCGACCTCGGCGGCGGAGCGCGCGC 882
QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
DB 883 CTGGCAGAGGTCGAAACAGTAACAAGAGCTGCCTCAGTCTACGATTTCTTTGATGGA 942
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
DB 943 ATCTATGCAAAATATCAGATGGTTATATACATCAGTGTGTGGCTCCAAATGTGAA 1002
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB 1003 GTACAGGTGAAAAATGGAGTATATATGAGAGAGTTTTTAAACTTACAGTCCGAGTGT 1062
QY 301 AspLeuValLeuAspAlaAlaHisGlyLysSerThrGluSerSerSerGlyProLysArg 320
DB 1063 GATTTGGTACTTGTATGCGCACATGAGAAAGATGACAGATCCAGTTCGGGGCCGAAACGT 1122
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys 340
DB 1123 GAAGAAATATATGGAGAGATTTTGTTCAAATGTTCAGACTTTGTGTGTGAGTTTAAA 1182
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB 1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAA 1242
QY 361 ValAsnGlyGluHisLysGlyLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla 380
DB 1243 GTGAATGGCGAAACAAAGAGAAGACCTGGAGCCCTGGATGTCAGGTGAATCAGACGCC 1302
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMet 400
DB 1303 AATGAGAACTTGAAGCTTTGAAAATGACGTATCTAATGATGGATGGATCCCAATGATATG 1362
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB 1363 TTTTCGATATATGAAGAAATATGTTGTTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB 1423 TATACAGTCCCTTAGAAGAGATAACTCAGAAGAAATTTTAAACGGGAAGCAAGGCA 1482
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1483 AACCAAGTTAGCAGAGAAATGAGTCAAGTGCCCGAGTACAAAGCTCGAGTGGCCCTGGAA 1542
QY 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB 1543 AATGATGATGAGGTGAGGAGAGAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGT 1602

QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
DB 1603 GAGGGGCACACATAAAACACTAGGAAATAAATATATTTCTCTCGGCAAAAGAAATAGA 1662
QY 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB 1663 GAAGTCATATCTCTGGGAGGTGGAGACAGAAATTCACCGCGTATGGGCAGCCTGGATCG 1722
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1723 GGTCTCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATTCCTGGTTCA 1782
QY 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
DB 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGGCATCGCCTTGCCCATCTCTTCTCTCT 1842
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
DB 1843 CGCCACCTTCTCGCTACAGTCCAGTCCCACTCTTCCACCTCGGCGCAGCCACCCCT 1902
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
DB 1903 ACACGGCGCGCTCCAGCGCCCTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
DB 1963 CATGGTTCTCCAGCTCTCTCTACTATGCTTAAACGATGTCTTTCAGAGAGGCGCTCCA 2022
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
DB 2023 AGGATGTCCCCAAGGCCCGCAGCATCTCTGAAATCACAGAGTTTCTGCTGGGAGGGT 2082
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
DB 2083 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACCCACCCAGTGAAGCAGCTACTCT 2142
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680
DB 2143 CCAGTAGCAAGACAGTCCCTCGGGGGAACGTGTCATCAGTGTGTCAGTGGGTTCCTCA 2202
QY 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
DB 2203 AGATTATCCCTTAAACTCATAGACCCAGGTCTCCAGACAGAACAGTATTGGAAATACC 2262
QY 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
DB 2263 CCCAGTGGCGCAGTCTTGTCTTCCCCAAGCTGCTATTATTCCAACTGAAGCTGTGCC 2322
QY 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGCTTATTCAGCTGTCATCTCTACGCTGTGTAGTCTGTGATCGAATCGAAGCTGTACC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2383 CTTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATCAGAGCAGAACTCTCTCCGAGG 2442
QY 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2443 AATAAAGAAATATTTAAACCCCAATGAAACATCACCTAGTCTTCAAAAGCTGAAACAAA 2502
QY 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
DB 2503 GGTATATCACCAGTTGTTCTGAAATAGAAACAGATTGATGATTTAAAGAAATTTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 AATGATTTTAGGTACAGCCCAAGTTCTACTTCTGAAATCTATGGATCAACTACTAACCAAA 2622
QY 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
DB 2623 AATAGAGAGGAGAAATCAAGAGATTTGATCAAAAGACAAAATTTGAACCAAGTCTAAG 2682
QY 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860

```
Db 2683 GATTCCTTTTCAATGAATAATAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAAATAGC 2742
Qy 861 ProSerIleSerProSerIleLeuSerAenThrGluHisLysArgGlyProGluValThr 880
Db 2743 CCAGCAATTCCTTCCATATCTAGTAACAGGAGCACAAGAGGGACCTGAGTCACT 2802
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900
Db 2803 TCCCAAGGGGTTTCAGACTTCAGCCCGCAGCATGTAAACAGAGAAAGACGATTAAGGAAGAG 2862
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 2863 AAGAAGAGCGCAGCTGAGCAAGTTAGGAATCAACATTAATCCCAATCAAGAGAGTTTC 2922
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2923 AACCCACGTTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCCTCAA 2982
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValThrThrGlnPro 960
Db 2983 GCACAACTAGCCCATCTATGTGTGGTGCATCAACAGCCAACTCCAGTTTATATCTCAGCCT 3042
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db 3043 GTTGTGTTTGCACCAATATGATGTATCCAGTCCCGAGTGGCCGCGGTGCAACCTTTTA 3102
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 3103 TACCCAAATACCTATGAGCCCATGCGCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCA 3162
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db 3163 AATATGCCCAACAGCGGCAAGCAGCAGCATCATCAGAGTGCCATGATGACCCAGCGTCA 3222
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 3223 GCAGCGGGGCCACCGATTGCGAGCCACCCACCGAGCTTACTCCAGCAATATGTGTGCTAC 3282
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 3283 AGTCCTCAGCAGGTTCCAAATCAGCCCTGTGTTGAGCATGTGCCACATATCATGCTCTCAG 3342
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCCTCATGTCTATAGTCTGTAAATCAGGGTAATGCTAGATGATGGCACCACCAACA 3402
Qy 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Db 3403 CACGCCCAGCCTGGTTTGTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACG 3462
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3463 CATGCGATGTATGATGTCCTCAATATACCATACACAGAGACAGCCCTCTTTCTTCTAC 3522
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTGCCATTTCCAGGGGCTCCCTTGTCTCAGCAGTATGGCCACCTTAACGCTACCTGTCAC 3582
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
Db 3583 CCACATACTCCACACCTCAGCCTTCAGTACCCCACTGGAGCAGCAGCAAGCCAAACAT 3642
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180
Db 3643 GGTGGAAGTATCTCTGACCCAGTCTCTTTCAGCACCCTCAGCACCAGGCGGCCAGGCT 3702
Qy 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3703 CTCCATCTGGCCAGTCACAGCAGCAGTACGCCATTTACACCGGGGCTTGGCCCAACT 3762
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
```

```
Db 3763 CCACCTCCATGACACACCTGCCTCCAAACAGCAGTGCACAGAGATAGTTTCCAGCAGCA 3822
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3823 CAACAGACTGTCTTTTACGATCCATCTCTCAGCTTCAGCGCGGTATACCAACCCACCC 3882
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3883 CACATGGGCCACGTACCTCAGGCTCATGTACAGTCAGGAATGGTCTCTTCATCCAAT 3942
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3943 GCCCATGGCCCAATCATGCTAATGACAGACACAGCCACCGCGGTCCCGAGCGGCCCTC 4002
Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
Db 4003 GCTCAAGTGCACACTACAGCCCATTCAGTCTCGACAACAGCGCATTTCCCTTATATGACG 4062
Qy 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db 4063 CACCTTCAGTACAAGCCACCCACCAACAGCAGTTG 4098

RESULT 6
ADZ49164
ID ADZ49164 standard; DNA; 4481 BP.
XX AC ADZ49164;
XX DT 30-JUN-2005 (first entry)
XX DE Insulin signaling pathway related gene, SEQ ID 493.
XX KW Antidiabetic; Gene Therapy; Non-insulin dependent diabetes;
XX KW insulin resistance; gene; ds.
XX OS Homo sapiens.
XX PN US2005085436-A1.
XX PD 21-APR-2005.
XX PF 08-JUL-2004; 2004US-00887553.
XX PR 08-JUL-2003; 2003US-0485883P.
XX PA (LIHH/) LI H.
XX PA (MAJU/) MA J.
XX PI Li H, Ma J;
XX WPI; 2005-305194/31.
DR Treating, preventing or ameliorating pathological conditions associated
XX with dysregulation of the insulin signaling pathway (ISP) comprises
XX administering to a subject an amount of a modulator of any of the
XX proteins regulated by ISP.
XX Disclosure; SEQ ID NO 493; 70pp; English.
XX The present invention relates to a method for treating, preventing or
XX ameliorating pathological conditions associated with dysregulation of the
XX insulin signaling pathway (ISP). The method comprises administering to a
XX subject a modulator for ISP-regulated proteins or a pharmaceutical
XX composition comprising the described modulator. The method is useful for
XX treating, preventing or ameliorating pathological conditions associated
XX with dysregulation of the ISP such as Type II diabetes or Type A syndrome
XX of insulin resistance. The present sequence is a human homolog of a
XX Drosophila gene regulated by ISP. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20050085436.
XX Sequence 4481 BP; 1144 A; 1380 C; 1014 G; 943 T; 0 U; 0 Other;
XX SQ
```

Alignment Scores:

Pred. No.: 5,276-200 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-802-228-2 (1-1312) x ADZ49164 (1-4481)

QY 1 MetArgSerAlaAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
DB 163 ATGCCTCAGCGCGCGAGCTCTCGAGTCCCGGCTGGCCACCGAGTCTCGCGCTTC 222
QY 21 AlaAlaAlaArgTTPProGlyTTPArgSerLeuGlnArgProAlaAlaArgArgSerGlyArg 40
DB 223 GCCGAGCCAGGTGGCCCGGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGGCGG 282
QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProProGlyPro 60
DB 283 GCGCGCGTGGCGCGCGCGCGGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCGCC 342
QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
DB 343 GCGCGCGCTCCCGCGCAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 402
QY 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyGlyPro 100
DB 403 GCGACGCGGGCGGCGGCTTTCGGCGCGGCTCCCGCGGCTCTTGGTCTCGCGGGCT 462
QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
DB 463 CCGCGCGCTTTCGTCGTCTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 522
QY 121 ProThrArgAlaSerProLeuGlyAlaAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB 523 CCAACCGCGCGCTCCCGCGCTCGCGCGCGCTCGCTCCCGCGCGCTCGCGCGCTCTCT 582
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB 583 GCGCGCGCGCTTCCCGCTGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCAC 642
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 643 ATGCTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762
QY 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
DB 763 AGCGGCTTCTAGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 822
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
DB 823 TCCTCGCGCACGGCTCCCTCCTCGGTGTTCGGGACCTCCGGCGCGCGCGCGCGCGCG 882
QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
DB 883 CTGGCAGAGGTCGAAACAGTAAACAGGACTGCGCTCAGTCTACGATTTCTTTTATGGA 942
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
DB 943 ATCTATGCAATATGAGGATGTTTCATATACTTACATCAGTTGTGGCTCCAAATGTGAA 1002
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB 1003 GTACAAGTGAAGATGGAGGTATATGAAGGAGTTTTTAAACCTTACAGTCGAAGTGT 1062
QY 301 AspLeuValLeuAspAlaAlaHisGlnLysSerThrGluSerSerSerGlyProLysArg 320

DB 1063 GATTTGGTACTTGATGCCGCACATGAGAAAAGTACAGAAATCCAGTTTCGGGGCCCCAACGT 1122
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys 340
DB 1123 GACAAATAATGGAGAGTATTTTCTCAAAATGTTTCAGACTTTGTGTGACAGTTTAA 1182
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaLysSerAlaLys 360
DB 1183 GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGTCTATCAGTCTAAA 1242
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTTPAspAlaGlyGluLeuThrAla 380
DB 1243 GTGAATGGCGAACAACAAGAGAGAGCTTGAGCGCTCGGATCGAGGTGAACACTCACGCC 1302
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTTPAspProAsnAspMet 400
DB 1303 AATCAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGATG 1362
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB 1363 TTTCCATATAATGAAGAAAAATATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB 1423 TATACAGTGCCTTTAGAAAGAGATACTCAGAAAGAAATTTTAAACCGGAAAGCAAGGCA 1482
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1483 AACCACTTAGCAGAAGAAATTTGAGTCAAGTCCCGCAGTACAAAGCTCGAGTGGCCCTGGA 1542
QY 461 AsnAspArgSerGluGluGlyTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB 1543 AATGATGATAGCAGTGAAGAAATAACACAGCAGTTACAGAGAAATTCAGTGAACGT 1602
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
DB 1603 GAGGGGCACAGCATAAACACATAGGGAATAATAATATATCTCTCTGGCAAGAAATAAGA 1662
QY 501 GluValIleSerTTPGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB 1663 GAAGTCATATCTCGGGAGTGGGAGACAGAAATTCACCGGTATGGGCGAGCTGGATCG 1722
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATCTGTGTCA 1782
QY 541 AspGlnArgValValAsnGlyValProTTPProSerProCysProSerProSerSerSer 560
DB 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGGCGCATCGCTTGCCTATCTCTCTCTCT 1842
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
DB 1843 CGCCACCTTCTCGCTACCGTCCAGTCCCAACTCTCTTCCACCTCGGCGACCCACCT 1902
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
DB 1903 ACAGCGCGCGCTCCAGCGCGCTCGCGGCGCATCCAGACCCCGCTCTCACCTCTCTCT 1962
QY 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
DB 1963 CATGGTCTCCAGTCTCTGTCTCTACTATGCTTAACGCATGTCTTTCAGAAAGGCGCTCCA 2022
QY 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
DB 2023 AGGATGTCCTCCAAAGGCCCGACATCTCGAAATCACAGAGTTTCTGTGGGAGGGGT 2082
QY 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
DB 2083 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACCCAGCCAGTGAAGCAGTACTCTCT 2142
QY 661 ProValAlaArgThrSerProSerGlyGlyThrTTPSerSerValValSerGlyValPro 680
DB 2143 CCAGTAGCAGGACAGTCCCTCGGGGGGAACGGGTGCATCAGTGTGTCAGTGGGTTCCTCA 2202

681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
 2203 AGATTATCCCTAAAACTCATAGACCCAGGTCTCCAGACAGAAAGTATTGGAATACC 2262
 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleLeuProThrGluAlaValAla 720
 2263 CCAGTGGGCCAGTCTCTGCTCTCCCAAGCTGGTATTATTCCAACTGAACTGTGGC 2322
 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
 2323 ATGCCTATTCCAGCTGCATCTCCTACGCTGCTAGTCTGCATCGAACAGAGCTGTACC 2382
 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
 2383 CCTTCTAGTGGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCGAAGTCTCTCGCAGG 2442
 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
 2443 AATAAAGAAAATATTAAACCCCAATGAACATCACTAGTCTTCTCAAAAAGCTGAAACAAA 2502
 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
 2503 GGTATATCACCACTGTTCTGACATAGAAAACAGATTGATGATTTAAAGAAATTTAAG 2562
 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
 2563 AATGATTTTAGGTTACAGCCAAAGTCTTCTCTGAATCTATGATCAACTACTATAACAAA 2622
 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
 2623 AATAGAGAGGGAGAAAATCAAGAGATTTGATCAAAAGACAAAATTTGAACCAAGTGTAAAG 2682
 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
 2683 GATTCCTTTCAATGAATAAGCAGCAGCAACTGTACCGAGTGGCAGCAGCAAGCCGAATAGC 2742
 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
 2743 CCCAGCATTTCCCTTTCAATCTTAGTAAACGAGAGCACAAGAGGGGACCTGAGGTCACT 2802
 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
 2803 TCCCAAGGGGTTCCAGCTTCCAGCCAGCATGTAAACAAGAGAAACGATAGGAAGAG 2862
 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
 2863 AAGAAAGACCCAGCTGAGCAGTGAAGAAATCAACATTGAATCCCAATGCAAGGAGTTC 2922
 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
 2923 AACCCAGTTCCCTTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCCTCA 2982
 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
 2983 GCACAACTAGCCCATCTATGGTGGGTGTCATCAACAGCCAACTCCAGTTTATCTCAGCCT 3042
 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
 3043 GTTTGTTTTCACCAAAATATGATGTATCCAGTCCAGTGAGCCAGCGGTGCAACCTTTA 3102
 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
 3103 TACCCATACCTATGACGCCCATGCCAGTGAATCAAGCCCAAGACATATAGAGCAGTACCA 3162
 1001 AsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
 3163 AATATGCCCAACAGCGCGCAAGACCAACATCATCAGAGTGCCATGATGCAACCCAGCGTCA 3222
 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
 3223 GCAGCGGGCCACCGATTGACAGCCACCCACCCAGCTTCTCCACCGCAATATGTTGCCTAC 3282

1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
 3283 AGTCTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAG 3342
 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
 3343 CATCTCATGTCTATAGTCTGTAAATACAGGTAATGTAGTAATGATGCGCACCAACA 3402
 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyValHisGluGlnThr 1100
 3403 CACGCCCCAGCTGGTTTGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462
 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
 3463 CATGGATGTATGATGTCATGTCCTCCCAATTACATACACAGGAGACAAGCCCTCTTCTTCTAC 3522
 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
 3523 TTTGCCATTTCACGGGCTCCCTTGTCTCAGCAGTATGGCACCTTAACGCTACCTGTGCAC 3582
 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
 3583 CCACATACTCCACACCTTCAGCTACCCCTCAGTACCCCTGAGCAGCAGCAAGCCAAACAT 3642
 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180
 3643 GGTGAAGTCTCTTCGACCCAGTCTGTTCAGCACCATCAGCACCAGCCGCCCAGGCT 3702
 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTCAGCCATTTTACCACGGGGCTTGCCTCAACT 3762
 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
 3763 CCACCTCTCCATGACACCTCTCCCTCCACAGCAGTCGCCACAGCAATAGTTTCCCGCAGCA 3822
 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
 3823 CAACAGACTGTCTTTACGATCCATCTCTCAGCTTACGGCGGTATACCAACCCACCC 3882
 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
 3883 CACATGGCCCGCTACCTCAGGCTCATGTACAGTCAGGAATGGTCTCTCTCTCATCCAAC 3942
 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
 3943 GCCCATGCGCCAAATGATGTAAATGACGACACAGCCACCCCGCGTCCCCAGCGCCCTC 4002
 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
 4003 GCTCAAGTGCATCAGGCCATTCCAGTCTCGAACACAGCGCATTTCCCTATATGACG 4062
 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
 4063 CACCTCTCAGTACAGCCCAACCCACCAACAGCAGTTG 4098
 RESULT 7
 AAV30270
 ID AAV30270 standard; DNA; 4367 BP.
 XX
 AC AAV30270;
 XX
 XX 02-OCT-1998 (first entry)
 XX
 XX Gene causative of spinocerebellar ataxia type 2 (SCA2) DNA sequence.
 XX
 XX Spinocerebellar ataxia type 2; SCA2; gene therapy; antisense therapy;
 XX CAG repeat; neurodegenerative disease; ds.
 XX
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX CDS 49..3990

```

FT      /*tag= a
FT      /product= "Spinocerebellar ataxia type 2 associated
FT      protein"
FT      repeat_region 544..612
FT      /*tag= b
FT      /*note= "normal CAG repeat region; this is increased in
FT      patients with SCA2"
FT      repeat_unit 544..546
FT      /*tag= c
FT      XX
FT      PN W09818920-A1.
FT      XX
FT      PD 07-MAY-1998.
FT      XX
FT      PF 30-OCT-1997; 97WO-JP003946.
FT      XX
FT      PR 30-OCT-1996; 96JP-00304059.
FT      XX
FT      PA (SRLS-) SRL INC.
FT      XX
FT      PI Tsuji S, Sanpei K;
FT      XX
FT      DR WPI; 1998-272215/24.
FT      XX
FT      DR P-PSDB; AAW60213.
FT      XX
FT      XX Nucleic acid fragments associated with spinocerebellar ataxia type 2 -
FT      PT contain increased number of CAG repeat region compared to normal gene.
FT      XX
FT      PS Claim 1; Page 13-22; 38pp; Japanese.
FT      XX
FT      CC This represents the sequence of a gene causative of spinocerebellar
FT      CC ataxia type 2 (SCA2), a neurodegenerative disease. This gene associated
FT      CC with SCA2, has a tri-nucleotide (CAG) repeat region which in the
FT      CC expression product produces a polyglutamine sequence from Gln-166 to Gln-
FT      CC 188. In the normal gene there are 15-25 CAG repeats but in SCA2 patients
FT      CC this number is increased to 35-100. Peptides encoded by nucleic acid
FT      CC fragments (DNA or RNA) containing sequences from the SCA2 associated
FT      CC gene, antibodies recognising the peptides and antisense nucleic acids
FT      CC hybridising with the nucleic acid fragments can be used for the
FT      CC investigation and diagnosis of SCA2. They can also be used for the
FT      CC treatment of SCA2 by antisense therapy or gene therapy
FT      XX
FT      SQ Sequence 4367 BP; 1124 A; 1328 C; 991 G; 924 T; 0 U; 0 Other;
FT      XX

Alignment Scores:
Pred. No.: 1,066-199 Length: 4367
Score: 6950.50 Matches: 1312
Percent Similarity: 99.92% Conservative: 0
Best Local Similarity: 99.92% Mismatches: 0
Query Match: 99.85% Indels: 1
DB: 2 Gaps: 1

US-10-802-228-2 (1-1312) x AAV30270 (1-4367)

QY      1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
DB      49 ATGCCTCAGCGCGCGAGCTCTCGAGTCCCGGCTCGCCAGCCAGCTCTCGCGCTTC 108
QY      21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSerGlyArg 40
DB      109 GCCGAGCCAGGTGCCCGGGGGCGCTCGCTCCAGCGCGCGCGCGCGAGCGGGCGG 168
QY      41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
DB      169 GSCGCGGTGGCGCGCGCGCGCGGACCGTATCCCTCCGCGCGCCCTCCCGCGCGCGCC 228
QY      61 GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
DB      229 GSCCGCGCTCCCTCCCGCGAGAGTCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 288
QY      81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyPro 100
DB      289 GGCAACGCGCGCGCGCGCTTTCGGCCCGCGCTCCCGCGCGCTCTCTGGTCTCGCGCGGCT 348

```

```

QY      101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
DB      349 CCCCGCCCTTTCGTCTGTCTCTCTCCCTCGCCAGCCCGGGCGCCCTCCGCGCGCG 408
QY      121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB      409 CCAACCCCGCGCTCCCGCTCGCGCGCGTGTCTCCCGCGCGCTTCCTCTCTG 468
QY      141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB      469 GCGCGCGCGGTCTCCCGGTGTCTCCCGCGCGGTGTCTGTGTATGGGCGCCCTCAC 528
QY      161 MetSerLeuLysPro---GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 179
DB      529 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 588
QY      180 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 199
DB      589 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 648
QY      200 GlySerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerValSer 219
DB      649 GCGAGCGCGCTTCTAGCTGCCCGCGCGCGCTTCCGCGCTCTCTCTCTCTCT 708
QY      220 SerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgPro 239
DB      709 TCGTCTCTCGCCACGGCTCTCTCTCGGTGTCTCGCGCGACCTCCGCGCGCGGAGGCC 768
QY      240 GlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAsp 259
DB      769 GGCCTTGGCAGAGGTGCAAAACAGTAACAAAGAGCTGCTCAGTCTACGATTCTTTGAT 828
QY      260 GlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCys 279
DB      829 GAATCTAATGAATAATAGGATGTTTATATATCTTACATCAGTTGTTGGCTTCAAAATG 888
QY      280 GluValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLys 299
DB      889 GAAGTACAAGTGAAATATGAGGTATATATGAGGAGTGTATTTTAACTTACAGTCCGAAG 948
QY      300 CysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLys 319
DB      949 TGTGATTTGGTACTTGTATGCGCGCATGAGAAAAGTACAGAAATCCAGTTCGGGGCGGAA 1008
QY      320 ArgGluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPhe 339
DB      1009 CGTGAAGAAATATAGGAGATATTTTGTTCAAATGTTTCAATGTTTGTGTGTGACAGTTT 1068
QY      340 LysAspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAla 359
DB      1069 AAAGATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCT 1128
QY      360 LysValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThr 379
DB      1129 AAAGTGAATGGCGAACAACAAAGAGAGAGACCTGGAGCCCTGGGATGCAAGTCACTACA 1188
QY      380 AlaAsnGluGluLeuLeuAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAsp 399
DB      1189 GCCAATGAGAACTTGGAGCTTTTGGAAATACAGCTATCTAATGGATGGATGCCCAATGAT 1248
QY      400 MetPheArgTyrAsnGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSer 419
DB      1249 ATGTTTCGATATATGAAGAAATATAGTGTAGTGTCTACGTATGATAGCAGTTTATCT 1308
QY      420 SerTyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArg 439
DB      1309 TCGTATACAGTGCCTTAGAAAAGAGATAACTCAGAAAGAAATTTTAAAAACGGGAAGCAAG 1368
QY      440 AlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeu 459
DB      1369 GCAAACCGATTAGCAGAAAGAAATTTGAGTCAAGTCCCGAGTCAAAAGCTCGAGTGGCCCTG 1428

```



```
Db      3589  GCTCTCCATCTGGCCAGTCCACAGCAGCAGTCAGCCATTTACACGCGGGCTTGGCCCA 3648
Qy      1200  ThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAla 1219
Db      3649  ACTCCACCCCTCATGACACACCTGCTCCAAACACGCGAGTCGCCACAGAAATAGTTTCCACGCA 3708
Qy      1220  AlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnPro 1239
Db      3709  GCACAACAGACTGTCTTTACCATCATCTTCTCAGCTTCAGCGCGGTATACCAACCCCA 3768
Qy      1240  ProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisPro 1259
Db      3769  CCCACATGGCCCACTACCTCAGGCTCATGTACAGTCAGGAATGTTCTTCTCATCCA 3828
Qy      1260  ThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAla 1279
Db      3829  ACTGCCATGGCCCAATGATGCTAATGACGACACAGCCACCGCGGTCCCGAGCGGCC 3888
Qy      1280  LeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMet 1299
Db      3889  CTCGCTCAAGTGTCACTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTATATG 3948
Qy      1300  ThrHisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312
Db      3949  ACGCACCTTCAGTACAGCCACCCACCACACAGCAGTTG 3987

RESULT 8
AAAT78912
ID      AAAT78912 standard; cDNA; 4200 BP.
XX
AC
XX
XX
DT      09-FEB-1998 (first entry)
XX
DE      Spinocerebellar ataxia gene SCA2.
XX
KW      Monoclonal antibody; neurodegenerative disease; polyglutamine; TBP;
KW      repeat region; affinity; TATA binding protein; Kennedy disease;
KW      transcription initiation factor; lymphoblastic cell line; schizophrenia;
KW      Huntington's disease; dominant autosomal spinocerebellar ataxia;
KW      x-linked spino-bulbar muscular atrophy; familial spastic paraplegia;
KW      dentatorubral-pallidolusial atrophy; bipolar affective disorder;
KW      manic depressive psychosis; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      misc_feature
FT      1. 3986 Location/Qualifiers
FT      /tag= g
FT      /note= "sequence contained in DAN1 clone"
FT      CDS
FT      3. .2747
FT      /tag= a
FT      /product= "SCA2 protein"
FT      /note= "this CDS contains a putative translational start
FT      codon for the SCA2 protein at positions 243-245"
FT      CDS
FT      3. .242
FT      /tag= c
FT      /note= "putative open reading frame which is in frame
FT      with the putative translational start site of the SCA2
FT      open reading frame"
FT      misc_signal
FT      239. .245
FT      /tag= d
FT      /note= "putative Kozak consensus signal"
FT      repeat_region
FT      258. .323
FT      /tag= e
FT      /note= "encodes polyglutamine repeat region; contains
FT      repeats of CAG with 2 CAA codons interspersed"
FT      repeat_unit
FT      258. .260
FT      /tag= f
FT      /note= "CAG repeats"
FT      CDS
FT      2594. .3640
FT      /tag= b
FT      /note= "this second open reading frame may be derived by
```

```
FT      misc_feature
FT      3987. .4200 a frameshift or by alternative splicing"
FT      /tag= h
FT      /note= "derived from the EST's AAH92640, AAN90240 and
FT      AA213574 from dbEST database"
FT      misc_feature
FT      4023. .4029
FT      /tag= i
FT      /note= "region which differs in length between the
FT      sequences of the EST clones AAH92640, AAN90240 and
FT      AA213574"
XX
XX      WO9717445-A1.
XX
XX      15-MAY-1997.
XX
XX      08-NOV-1996; 96WO-FR001773.
XX
XX      10-NOV-1995; 95FR-00013576.
XX      (CNRS ) CNRS CENT NAT RECH SCI.
XX      (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
XX      Tora L, Lutz Y, Trottier Y, Mandel J;
XX
XX      WPI; 1997-281034/25.
XX      P-PSDB; AAW24800, AAW24801.
XX
XX      Antibody 1C2 used for treating or preventing neuro-degenerative diseases
XX      - associated with proteins containing long poly:glutamine repeats, e.g.
XX      Huntington's disease.
XX
XX      Claim 21; Page 45-47; 69pp; French.
XX
XX      The invention relates to a monoclonal antibody (Mab) 1C2 for the
XX      treatment of neurodegenerative diseases associated with the presence of
XX      polyglutamine repeat regions. This Mab is already known for its affinity
XX      to the TATA binding protein (TBP) transcription initiation factor,
XX      especially at the amino acid sequence LEEQQRQQQQQ found at the N-
XX      terminus of TBP. Mab 1C2 has been shown to have a high affinity for
XX      polyglutamine repeats with a proportional affinity to the number of
XX      glutamine repeats. This affinity has been used to identify genes encoding
XX      proteins containing long polyglutamine repeats which are implicated in
XX      neurodegenerative diseases. A screen of an expression library, generated
XX      from a lymphoblastic cell line from a patient suffering from
XX      spinocerebellar ataxia (SCA), with Mab 1C2 isolated 6 new sequences
XX      (AAAT78906-T78911) encoding polyglutamine repeats. Mab 1C2 also isolated
XX      the complete SCA2 gene in clone DAN1 (sequence presented here). The
XX      sequence appears to contain 2 open reading frames (ORF) the second of
XX      which may be generated by an frameshift slippage or by an alternative
XX      splicing event. The first ORF also encodes a 22 amino acid polyglutamine
XX      repeat region near the N-terminus of the protein. Normal SCA2 alleles
XX      contain 17-29 CAG triplet repeats with 1-3 CAA repeats interspersed
XX      whereas the mutant sequence from patients with SCA contains at least 10,
XX      preferably 37-50 CAG repeats. Mab 1C2, active fragment of it or nucleic
XX      acids encoding it are specifically used to treat Huntington's disease,
XX      SCA types 1-5 or 7, X-linked spino-bulbar muscular atrophy (Kennedy
XX      disease), dentatorubral-pallidolusial atrophy, dominant autosomal
XX      spinocerebellar ataxia, familial spastic paraplegia, bipolar affective
XX      disorder, manic depressive psychoses and schizophrenia
XX
XX      Sequence 4200 BP; 1152 A; 1200 C; 913 G; 935 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 1.49e-182 Length: 4200
XX      Score: 6377.00 Matches: 1211
XX      Percent Similarity: 98.30% Conservative: 0
XX      Best Local Similarity: 98.30% Mismatches: 1
XX      Query Match: 91.61% Indels: 21
XX      DB: 2 Gaps: 1
XX
XX      US-10-802-228-2 (1-1312) x AAAT78912 (1-4200)
XX
XX      Qy 81 GlyAsnGlyGlyAlaPheArgProGlySerArgLeuLeuGlyGlyPro 100
```

Db 3 GGCACCGCGCGCGCGCTTCGCGCGCGCTCCCGCGCGCTCCCTGGTCTCGCGCGCGCT 62
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
Db 63 CCCCAGCCCTTCGTCT 122
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 123 CCAACCGCGCGCTCCCGCGCGCGCTCGCGCGCGCTCCCGCGCGCTCCCTCTCTCTG 182
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 183 GCGCGCGCGCTCCCGCGCGCTCTCCCGCGCGCGCTCGAGCGGTGTATGGCGCGCTCAC 242
Qy 161 MetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 243 ATGTCTGAGACCCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 302
Qy 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 303 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 362
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerSerSerSer 220
Db 363 AGCGCGCTTCTAGCGCTCCCGCGCGCGCTTCGCGCGCTTCGCGCTCTCTCTCTG 422
Qy 221 SerSerAlaThrAlaProSerValValAlaAlaThrSerGlyGlyGlyGlyGlyGly 240
Db 423 TCCTCGCCACCGCTCCCT 482
Qy 241 LeuGlyArgGlyArgAsnSerAsnLeuGlyLeuProGlnSerThrIleSerPheAspGly 260
Db 483 CTGGGCAGAGGTCGAAACAGTAAACAGAGTGCCTCAGCTACGATTTCTTTTGTATGA 542
Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
Db 543 ATCTATGCAAAATATGAGGATGTTTATATATATATATATATATATATATATATAT 602
Qy 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db 603 GTACAGTGAAATATGAGGATATATATATATATATATATATATATATATATATATAT 662
Qy 301 AspLeuValLeuAspAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
Db 663 GATTTGGTACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 722
Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValValGlnPheLys 340
Db 723 GAAGAAATATGAGAGATATTTTGTTCAAATGTTTCAGACTTTTGTGTGTATACAGTTTAA 782
Qy 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db 783 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTCTCTATCAGTGTCTAAA 842
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380
Db 843 GTGATGGCGCAACACAAAGAGAGGACCTGGAGCCCTGGGATGTCAGTGAACCTCACAGCC 902
Qy 381 AsnGluGluLeuAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
Db 903 AATGAGAACTTGAGGCTTTGGAAATGACGTATCTAATGATGGATGCCAATGATATG 962
Qy 401 PheArgTyrAsnGluLeuThrGlyValValValSerThrTyrAspSerSerLeuSerSer 420
Db 963 TTTTCGATATATAGAGAAATATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1022
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db 1023 TATACAGTCCCTTAGAAGAGATATCTCAGAGAAATTTTAAACCGGAGCAGCGCA 1082
Qy 441 AsnGlnLeuAlaGluLeuGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460

Db 1083 AACCAAGTTAGCAGAGAAATTTGAGTCAAGTGCCCAAGTACAAAGCTCGAGTGGCCCTGGAA 1142
Qy 461 AsnAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db 1143 AATGATGATAGAGTGAG 1202
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGluGlyGlnArgAsnArg 500
Db 1203 GAGGGGCACAGCATAAACACTAGGAGAAATAAATATATATTTCTCTCGGACAAAGAAATAGA 1262
Qy 501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1263 GAAGTCATATCTGGGGAAGTGGGAGAGAGAGATTCACCGCGTATGGCCAGCCTCGATCG 1322
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1323 GGTCTCATGTCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATTTCTGGTTCA 1382
Qy 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560
Db 1383 GACCAAGAGTAGTTAATGGAGGTGTTCCTGGCCATCGCCTTGGCCATCTCTCTCTCTCT 1442
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1443 CCCCCACCTTCTCGCTACAGTCAGTCCCAACTCTCTTCCACCTCGGCGCAGCCACCT 1502
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db 1503 ACACGGCGCCCTCCAGGCCCTCCGGGCCATCCAGACCCCGCTCTCACCTCTCTCTCT 1562
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1563 CATGGTTCTCCAGCTCCTGTCTCTATATGCTTAAACGCATGTCTTCAAGAGGGCTCCA 1622
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Db 1623 AGGATGTCCCCAAAGGCCAGGACATCTCTCGAAATCACAGAGTTTCTCTCTGGAGGGT 1682
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
Db 1683 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACACCCAGCCAGTGAAGCAGTACTCTCT 1742
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTyrSerSerValValSerGlyValPro 680
Db 1743 CCAGTAGCAAGACCCAGTCCCTCGGGGGAAGCTGTGTATCATGTGTCTAGTGGGGTTCCA 1802
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db 1803 AGATTATCCCTTAAACACTATAGACCCAGGTCTCCAGACAGAACAGTATTTGAAATACC 1862
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
Db 1863 CCCAGTGGGCCAGTCTTCTCTCCCAAGCTGGTATTTATTTCAACTGAAGCTGTGTGCC 1922
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
Db 1923 ATGCTATTTCCAGCTGTCATCTCTACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1982
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
Db 1983 CCTTCTAGTAGGCTTAAAGATTCAGGCTTCAAGATCAGAGGAGAACCTCTCTCTCGAGGG 2042
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
Db 2043 AATAAGAGAAATATTAACCCCAATGAACATCACCTAGTCTCTCAAAAGCTGAACACAA 2102
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
Db 2103 GGTATATCACAGTGTCTTCTGAACATAGAAACAGATTTGATGATTTAAAGAAATTTAAG 2162
Qy 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db 2163 AATGATTTTAGGTTACAGCCAAAGTTCTACTTCTGAAATCTATGGATCAACTACTTAAACAA 2222

```
Qy 821 AsnArgGluGlyGluLysSerArgAspLeuLeuIleLysAspLysIleGluProSerAlaLys 840
Db 2223 AATAGAGAGGGAGAAAATCAAGAGATTGATCAAGAGACAAAATTTGAACCAAGTGCTAAG 2282
Qy 841 AspSerPheIleGluAenSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Db 2283 GATTCTTTTCATTGAAATPAGCAGCAGCAACTGTACACAGTGCAGCAGCAAGCCGAATAGC 2342
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db 2343 CCCAGCATTTCCCTTCAATTAATCTAGTAACAGGAGCAAGAGGGACCTGAGGTCACT 2402
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
Db 2403 TCCCAAGGGGTTTCAGACTTCAGGCCAGCATGTTAAACAAGAGAAAGACGATAAGAGAG 2462
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAenProAsnAlaLysGluPhe 920
Db 2463 AAGAAAGACGCGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCCAAGGAGTTC 2522
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2523 AACCCACGTTCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCCTCA 2582
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db 2583 GCACAACCTAGCCCATCTATGTTGGGTTCATCAACAGGCCAACTCCAGATTTTACTCAGCCT 2642
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db 2643 GTTTGTTTTCACCAAAATATGATGATATCCAGTCCAGTGAGCCCGAGCGGTGCAATA---- 2698
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 2699 -----CCCA 2701
Qy 1001 AsnMetProGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db 2702 AATATGCCCAACAGCGGCAAGACCAGCATCATCAGAGTGCCATGATGACCCAGCGTCA 2761
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 2762 GCAGCGGGCCACCGATTGCGAGCCACCCAGCAGCTTACTCCAGCAGATATGTTGCGCTAC 2821
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 2822 AGTCTCAGCAGTTCCTCCAAATCAGCCCTTGTTCAGCATGTGCCACATATTCAGTCTCAG 2881
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 2882 CATCTCATGTCTATAGTCTGTAATACAGGGTAATGCTAGATGATGGCACCAACCA 2941
Qy 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Db 2942 CACGCCAGCGCTGGTTAGTATCTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGC 3001
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3002 CATGGATGATGATGATGCTCCCAATATACCATACAAAGGAGACAGCCCTTCTTCTAC 3061
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3062 TTTGCCATTTCCAGGGGTCCTTGTCTCAGCAGTATGGCACCTAACGCTACCTCGCAC 3121
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db 3122 CCACATACTCCACACCTTCAGCTTTCAGCTTACCCCTCAGTGCAGCAGCAAGCAACAT 3181
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGlnAla 1180
Db 3182 GGTGGAGTCTATCTGCACCCAGTCTCTGTTGAGCACCNCATCAGCAGCCGCCGCCAGGCT 3241
```

```
Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3242 CTCATCTGGCCAGTCCACAGCAGCAGTCAGCCATTTACCACGCGGGCTTGGCCAACT 3301
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3302 CCACCTCCATGACACCTGCCCTCCAACACGAGTGCACAGAAATAGTTTCCAGCAGCA 3361
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3362 CAACAGACTGTCTTTAGATCCATCTCTCAGCTTACGCCGGCTATACCAACCCACCC 3421
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3422 CACATGCCCACTGATACCTCAGGCTCATGTACAGTCAGAAATGGTTCTCTTCATCCAAC 3481
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3482 GCCCATGGCCAATGATGCTAATACGACACAGCAGCCGCGGTCCCCAGGCGCCCTC 3541
Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
Db 3542 GCTCAAGTGCATCAGACCCCATTTCCAGTCTCGAACACAGCGCATTTCCCTATATGAGC 3601
Qy 1301 HisProSerValGlnAlaHisHisGlnGlnLeu 1312
Db 3602 CACCTTCAGTACAAAGCCACCAACAGCAGTTG 3637
RESULT 9
ABV75290
ID ABV75290 standard; DNA; 4225 BP.
XX
AC ABV75290;
XX
DT 07-MAR-2003 (first entry)
XX
DE Mouse SCA2 polypeptide encoding DNA.
XX
KW SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity;
XX memory impairment; mutant; mouse; gene; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
CDS 27..3884
FT /*tag= a
FT /transl_except= (pos: 222..224, aa: Xaa)
FT /note= "Xaa is Pro"
FT /product= "SCA2"
XX
WO200289567-A2.
XX
PD 14-NOV-2002.
XX
PF 07-MAY-2002; 2002WO-US014698.
XX
PR 07-MAY-2001; 2001US-0289231P.
XX
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI Pulst SM;
XX
DR WPI; 2003-103467/09.
DR P-PSDB; ABB82697.
XX
PT New mutant mouse, useful as a disease model for identifying agents for
PT treating obesity or memory impairment, comprises a disrupted
PT spinocerebellar ataxia 2 gene.
XX
PS Disclosure; Page 56-62; 76pp; English.
XX
CC The invention relates to a mutant non-human mammal having a disrupted
CC spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a
```


XX Mouse SCA2 cDNA encoding ataxin-2.
 KW SCA2 gene; spinocerebellar ataxia-2; ataxin-2; mouse; diagnosis;
 KW olivo-ponto-cerebellar atrophy; ss; ds.
 XX Mus musculus.
 OS
 FH Key Location/Qualifiers
 FT CDS 50..3457
 FT /*tag= a
 XX
 PN W09742314-Al.
 XX
 PD 13-NOV-1997.
 XX
 PF 08-MAY-1997; 97WO-US007725.
 XX
 PR 08-MAY-1996; 96US-0017388P.
 PR 19-JUL-1996; 96US-0022207P.
 PR 08-OCT-1996; 96US-00727084.
 XX
 XX (CEDA-) CEDARS SINAI MEDICAL CENT.
 PA
 XX
 PI Pulst S;
 XX
 DR WPI; 1998-086523/08.
 DR P-PSDB; AAW33808.
 XX
 XX Nucleic acids encoding human and mouse ataxin 2 - a product of the
 FT spinocerebellar ataxia 2 gene, SCA2; useful in the diagnosis of ataxia
 PT type 2.
 XX
 PS Claim 6; Page 63-69; 98pp; English.
 XX
 CC This cDNA sequence corresponds to a novel SCA2 gene encoding a mouse
 CC spinocerebellar ataxin-2 (SCA2) polypeptide, designated ataxin-2 (see
 CC AAW33808). It was isolated from a newborn mouse brain cDNA library using
 CC human SCA2 cDNA (see AAV06552) as probe. Ataxia type 2 can be diagnosed
 CC by detecting a genomic or transcribed mRNA sequence in an individual
 CC having an expanded CAG repeat corresponding to the CAG repeat region of
 CC the human SCA2 gene. Also claimed are kits for detecting mutations at the
 CC SCA2 locus, antisense oligonucleotides, and transgenic animals useful for
 CC studying the physiological roles of ataxin-2 and its effect upon
 CC behaviour
 XX
 SQ Sequence 3798 BP; 992 A; 1118 C; 891 G; 797 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,566-159 Length: 3798
 Score: 5606.50 Matches: 1063
 Percent Similarity: 94.20% Conservative: 41
 Best Local Similarity: 90.70% Mismatches: 37
 Query Match: 80.54% Indels: 31
 DB: 2 Gaps: 8
 US-10-802-228-2 (1-1312) x AAV06553 (1-3798)
 QY 148 ProArgProAlaCysGluProValTyrGlyProLeuThrMetSerLeuLysProGlnGln 167
 DB 11 CCGCGCGCGCGTGCAGCGCGGTGTATGGCGCGCTACCATGTGCTGTGAAGCCGCGCG 70
 QY 168 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 187
 DB 71 CAG-----
 QY 188 ProProProAlaAlaAlaAsnValArgLysProGlySerGlyLeuLeuAlaSerPro 207
 DB 74 ---CG 127
 QY 208 AlaAlaAlaProSerProSerProSerSerValSerSerSerAlaThrAlaProSer 227
 DB 128 GCG 184

QY 228 SerValValAlaAlaThrSer-----GlyGlyGlyArgProGlyLeuGlyArgGly 244
 DB :::::|||||
 QY 185 CGCGCGGTGGGTCTTCTCG 244
 DB :::::|||||
 QY 245 ArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGlyIleTyrAlaAsn 264
 DB :::::|||||
 QY 245 CGGAACAGTAGCAAGGAGCTGCCTCAGCTACGATTTCTTTTGATGGAATCTATGCAAC 304
 DB :::::|||||
 QY 265 MetArgMetValHisIleLeuThrSerValValGlySerLysCysGluValGlnValLys 284
 DB :::::|||||
 QY 305 GTGAGGATGGTTCTATACCTACGTTCAGTTGTGGATCGAAATGTCAAGTACAAAGTAA 364
 DB :::::|||||
 QY 285 AsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCysAspLeuValLeu 304
 DB :::::|||||
 QY 365 AACGGAGGCATATATGAAGGAGTTTTTAAACATACAGTCTTAAGTGTGACTTGGTACTT 424
 DB :::::|||||
 QY 305 AspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArgGluGluIleMet 324
 DB :::::|||||
 QY 425 GATGCTGCACATGAGAAAGTACAGATCCAGTTCCGGGGCCAAAGCTCAAGAAATAATG 484
 DB :::::|||||
 QY 325 GluSerIleLeuPheLysCysSerAspPheValValGlnPheLysAspMetAspSer 344
 DB :::::|||||
 QY 485 GAGAGTGTTCCTCAATGCTCAGACTTCGTTCGTGTACAGTTTAAAGATACAGACTCC 544
 DB :::::|||||
 QY 345 SerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLysValAsnGlyGlu 364
 DB :::::|||||
 QY 545 AGTTATGCACGAGAGATGCTTTTACTGACTCTGCTCTCAGCGCAAGGTGAATGGTGAG 604
 DB :::::|||||
 QY 365 HisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAlaAsnGluGluLeu 384
 DB :::::|||||
 QY 605 CACAGAGGAAGGACCTGGAGCCCTGGAGCTCAGGGGAGCTCAGGCCAGCGAGGAGCTG 664
 DB :::::|||||
 QY 385 GluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMetPheArgTyrAsn 404
 DB :::::|||||
 QY 665 GAG---CTGAGAAATGATGTGTCTAATGATGGACCCCAATGACATGTTTCGATATAAT 721
 DB :::::|||||
 QY 405 GluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSerTyrThrValPro 424
 DB :::::|||||
 QY 722 GAAGAGAATTATGGTGTGTGTCCACATATGATAGCAGTTTATCTTCATATACGGTTCT 781
 DB :::::|||||
 QY 425 LeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAlaAsnGlnLeuAla 444
 DB :::::|||||
 QY 782 TTAGAAGGAGCAACTCAGAAAGATTTCTTAACGGGAGGCAAGGCAACCAAGTTAGCA 841
 DB :::::|||||
 QY 445 GluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGluAsnAspAspArg 464
 DB :::::|||||
 QY 842 GAAGAAATTTGAATCCAGTGTCTCAGTACAAAGCTCGTGTGCGCCCTTGAGAATGATGACCG 901
 DB :::::|||||
 QY 465 SerGluGluGlyLysTyrThrAlaValGlnArgAsnSerSerGluArgGluGlyHisSer 484
 DB :::::|||||
 QY 902 AGTGAGGAAGAAAAATACACAGCAGTCCAGAGAACTGCAGTGCACCGGAGGGGCATGCG 961
 DB :::::|||||
 QY 485 IleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArgGluValIleSer 504
 DB :::::|||||
 QY 962 CCCAACACTAGGACAAATAATATATTCCTCTCGACAAAGAAACAGAGAGTCTCTATCC 1021
 DB :::::|||||
 QY 505 TrpGlySerGlyValArgGlnAsnSerProArgMetGlyGlnProGlySerGlySerMetPro 524
 DB :::::|||||
 QY 1022 TGGGGAAGTGGGAGACAGAGCTCACCGAGTGGGCGAGCTCGGGCCAGGCTCCATGCCG 1081
 DB :::::|||||
 QY 525 SerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySerAspGlnArgVal 544
 DB :::::|||||
 QY 1082 TCAAGAGCTCTCTCACACTTCAGATTTCAACCCGAACTGGCTGGCTCAGACCAAGAGTA 1141
 DB :::::|||||
 QY 545 ValAsnGlyGlyValProTyrProSerProCysProSerProSerSerArgProProSer 564
 DB :::::|||||
 QY 1142 GTTAATGAGGAGTTCCTCGGCGCATCGCTTGCCTTCTCATTTCTCTCGCCACCTTCT 1201
 DB :::::|||||
 QY 565 ArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrProThrArgProPro 584
 DB :::::|||||
 QY 1202 CGCTACAGTCCAGTCCCAACTCTCTTCCACCTCGGCGACCCACCCATACACGCGCCGCC 1261
 DB :::::|||||


```
QY 585 SerArgProSerArgProSerArgProProSerHisProSerAlaHisGlySerPro 604
DB 1262 TCCAGGCCCTCCAGGCCATCCAGACCCCGTCTCACCCCTCTGCTCATGGTTCTCCA 1321
QY 605 AlaProValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSerPro 624
DB 1322 GCTCTGTCTCTACTATGCTTAAAGCATGTCTTCAGAAAGNACCCCAAGGATGCTCCA 1381
QY 625 LysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSer 644
DB 1382 AAGGCACAGCGCCACCCCTCGGAATCACAGAGTCTCTGCTGGAGAGGCTCCCATGTCTAGT 1441
QY 645 GlyLeuGlnPheValSerHisAsnProProSerGluAlaAlaThrProProValAlaArg 664
DB 1442 GGCCTAGAAATTTGTATCCCAATCCCAATCCCAATGGAAGCAGCTGCTCTCCAGTGGCAAGG 1501
QY 665 ThrSerProSerGlyGlyThrTriSerSerValValSerGlyValProArgLeuSerPro 684
DB 1502 ACCAGTCTCGAGGGGAAGGTGTCCTCAGTGTCTCAGTGGGTTCGAGGTTTCAAGGTTATCTCCC 1561
QY 685 LysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyPro 704
DB 1562 AAAACTCACAGACCCAGGCTCTCCAGGCAGAGCAGCATTGGAAACTCTCCAGCGGGGCT 1621
QY 705 ValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAlaMetProIlePro 724
DB 1622 GTGCTGTGTTCTCCCAAGCTGGCATCTCCTCAGTGTCTCAGTGGGTTCGAGGCGGTTCCTCCAGTGGTCTCC 1681
QY 725 AlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGlu 744
DB 1682 GCGCATCTCCGACTCTCTCCAGCCCTGCATCCACAGAGCAGTACCCCATCTATTGAG 1741
QY 745 AlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsn 764
DB 1742 GCAAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGCAGGAGTAAAGAAAT 1801
QY 765 IleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerPro 784
DB 1802 GTTAAAGCAAGTGAACATCACTAGCTTTTCAAAGCTGACAAAGATGATGTGCCA 1861
QY 785 ValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsnAspPheArg 804
DB 1862 GTTGTCTCTGAACACAGAAAACAGATTGATCTTAAAGAGTTTAAAGATGATTTTAGG 1921
QY 805 LeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGly 824
DB 1922 TTACAGCCAAGCTCTACATCTGAATCTATGGATCAACTACTTAAGCAAAATAGAGAAGG 1981
QY 825 GluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysAspSerPheIle 844
DB 1982 GAAAAGTCACGAGATTTTGATTTAAAGATAAAAACGGAAGCAAGTGCTTAAGGATAGTTTCATT 2041
QY 845 Glu-----AsnSerSerAsnCystrSerGlySerSerLysProAsnSerProSer 862
DB 2042 GACAGCAGCAGCAGCAGCAGCACTGTACAGTGGCAGCAGCAGCAAGCAACAGCCCTAGC 2101
QY 863 IleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThrSerGln 882
DB 2102 ATCTCCCTTCCATGCTTAGTAAATGCAGAGCACAAGAGGGGCGCTGAGGTACATCCCAA 2161
QY 883 GlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluLysLys 902
DB 2162 GGGGTGCAGATCTTCAGGCCAGCTTGCACAAAGAGAGAGGATGACAGAGAAGAGAGAA 2221
QY 903 AspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPheAsnPro 922
DB 2222 GACACAAACAGAGCAGGTTAGGAAATCGACATTGAATCCCAATGCAAGAGGTTCAACCCCT 2281
QY 923 ArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGlnAlaGln 942
DB 2282 CGTTCTTTCTCTCAGCAAAAGCCTTCTACTACCCCAACGTCACCTCGGCGCTCAAGACAA 2341
QY 943 ProSerProSerMetValGlyHisGlnGlnProThrProValThrThrGlnProValCys 962
```

DB:		DB:		DB:	
US-10-802-228-2	(1-1312) x AAF21628	(1-2006)	Gaps:	1	
QY	850 AenCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSer	869			
DB	6 AGCTGTACCATGGCGAGCCGAATAGCCCCAGCATTTTCCCCTTCATACTTAGT	65			
QY	870 AenThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerPro	889			
DB	66 AACACGAGCACAGAGGGGCCTTGAGGTCACTTCCCAAGGGGTTTCAGACTTCCAGCCCA	125			
QY	890 AlaCysLysGlnGluLysAspLysGluGluLysLysAspAlaAalaGluGlnValArg	909			
DB	126 GCATTAAACAAGAGAAGACCATTAAGGAAGAGAAGAACGACGCTGAGCAAGTTAGG	185			
QY	910 LysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProLys	929			
DB	186 AAATCAAATTAATCCCAATGCCAAGGAGTTCAACCCACGCTTCTCTCAGCAAAAG	245			
QY	930 ProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetValGly	949			
DB	246 CCTTCTACTACCCCAACTTCACTCGGCCTCAAGCACCACTTAGCCCATCTATGGTGGT	305			
QY	950 HisGlnGlnProThrProValThrGlnProValCysPheAlaProAsnMetMetTyr	969			
DB	306 CATCAACAGCCCACTCCAGTTTTATCTCAGCCTGTTTTTTTGGACCACCAATATGATGAT	365			
QY	970 ProValProValSerProGlyValGlnProLeuTyrProIleProMetThrProMetPro	989			
DB	366 CCAGTCCCAGTGAGCCAGCGCTGCAACCTTTATACCCAATACCTATAGAGCCCATGCCA	425			
QY	990 ValAsnGlnAlaLysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGln	1007			
DB	426 GTGAATCAAGCCCAAGACATATAGAGCAGCGTAAAGTACCAATATATGCCCAACAGCGCAA	485			
QY	1008 AspGlnHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAla	1027			
DB	486 GACCAAGCATCATCAGAGTGCCATGATGCCAGCGGTGACAGCGGGCCACCATGTCGA	545			
QY	1028 AlaThrProProAlaThrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsn	1047			
DB	546 GMAACCCCAAGCTTACTCCACGCNAATATGTTGCTACAGTCTCAGCAGTTCCTCCAAAT	605			
QY	1048 GlnProLeuValGlnHisValProHisTyrGlnSerGlnHisHiproHisValTyrSerPro	1067			
DB	606 CAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGTCTCT	665			
QY	1068 ValileGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuVal	1087			
DB	666 GTAATACAGGGTAATGCTAGAATGATGGCACCAACCAACACCGCCAGCGCTGGTTAGTA	725			
QY	1088 SerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysPro	1107			
DB	726 TCTTCTTACGAACTCAGTACGGGGCTCATGACAGCAGCATGCGATGATGATGATGTC	785			
QY	1108 LysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySer	1127			
DB	786 AAATTACCATACAACAGGAGACAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	845			
QY	1128 LeuAlaGlnGlnTyr-AlaHisProAsnAlaThrLeuHisProHisThrProHisProGl	1147			
DB	846 CTTCGCTCAGCAGTATNGCRACCTTAACGCTACCTTGCACCCACATCTCCACACCTCA	905			
QY	1147 nProSerAlaThrProThrGlyGlnGlnGlnSerGlnHisGlyGlySerHisProAlaPr	1167			
DB	906 GCCTTCAGCTTACCCCTCAGCAGGACCAACCAACATGGTGGAGTGCATCTCTGCACC	965			
QY	1167 oSerProValGlnHisGlnHisGlnAlaAlaGlnAla-LeuHisLeuAlaSerProG	1187			
DB	966 CAGTCTCTGTTACGACCATCATGACAGCGCGCCAGGCTNCTCCATCTCGGCAGTCCAC	1025			
QY	1187 lnglnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProA	1207			

DB: 3419 CACCCCTTCAGTCAAGCCCAACCAACCAAGCAGCTTG 3454

RESULT 11
AAF21628
ID AAF21628 standard; DNA; 2006 BP.
XX AC AAF21628;
XX DT 27-MAR-2001 (first entry)
XX DE Human breast and ovarian cancer associated antigen gene SEQ ID 15.
XX KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW antitumor; neuroprotective; antiviral; anti-allergic; hepatotropic;
KW antidiabetic; anti-inflammatory; anticancer; vulnarary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.
OS Homo sapiens.
XX PN W0200055173-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WO-US005881.
XX PR 12-MAR-1999; 99US-0124270P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Ruben SM;
XX WPI; 2000-611515/58.
DR P-PSDB; AAB58725.
XX New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention, treatment
PT and diagnosis of cancer, immune disorders, cardiovascular disorders and
PT neurological diseases.
XX Claim 1; Page 490-491; 1299pp; English.
XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive; neutropic;
CC neuroprotective; antiviral; anti-allergic; hepatotropic; antidiabetic;
CC antiinflammatory; anticancer; vulnarary; anticonvulsant; antibacterial;
CC antifungal; antiparasitic and cardiant activity. The polynucleotide and
CC protein sequences are used in the diagnosis of cancer, particularly
CC breast and ovarian cancer. The nucleic acid sequences, proteins, agonists
CC and agonists may also be used in the diagnosis, prevention and treatment
CC of immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC cardiovascular disorders such as myocardial ischaemias; wound healing;
CC neurological diseases such as cerebral anoxia and epilepsy; and
CC infectious diseases
XX Sequence 2006 BP; 598 A; 577 C; 370 G; 456 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 1,09e-65 Length: 2006
Score: 2463.00 Matches: 460
Percent Similarity: 98.72% Conservative: 1
Best Local Similarity: 98.50% Mismatches: 2
Query Match: 35.38% Indels: 4

Db 1026 AGCAGCAGTCACCAATTTACACGCGGGCTTGCCTCCCACTCCACCTCCATGACACCTG 1085
QY 1207 laSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrI 1227
Db 1086 CCTCCAAACAGCGCAGTCGCCACAGATAGTTTCCACAGCAGCACAACAGACTGTTTACGA 1145
QY 1227 leHisProSerHisValGlnProAlaThrProAlaThrAsnProHisMetAlaHisValProG 1247
Db 1146 TCCATTCCTTCTCACGTTTCAGCGGGGTATACCAACCCACCCACATGGCCACGTCACCTC 1205
QY 1247 lnAlaHisValGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetL 1267
Db 1206 AGGCTCATGTACAGTCAGAAATGGTTCCTTCATCCAACTGCCCATGCCCAATGATGC 1265
QY 1267 euMetThrThrGlnProProGlyGlyProGlnAlaAlaAlaGlnSerAlaLeuGlnP 1287
Db 1266 TAATGACGACACAGCCACCCGCGGTCCCGAGCGCCCTCGCTCAAAGTGCACCTACAGC 1325
QY 1287 rolleProValSerThrThrAlaHisPheProTyMetThrHisProSerValGlnAlaH 1307
Db 1326 CCATTCAGTCTCGAACAAGCGCATTTCCCTATATGACGACCCCTTCAGTACAAGCCC 1385
QY 1307 leHisGlnGlnGlnLeu 1312
Db 1386 ACCACCAACAGCAGTTG 1402

RESULT 12

ABV75289
ID ABV75289 standard; DNA; 1257 BP.
XX
AC ABV75289;
XX
DT 07-MAR-2003 (first entry)
XX
DE Mouse SCA2 partial polypeptide encoding DNA.
XX
KW SCA2; spinocerebellar ataxia 2; anorectic; nootropic; obesity;
KW memory impairment; mutant; mouse; gene; ds.
XX
OS Mus musculus.
FH Key Location/Qualifiers
FT CDS 2..1255
FT FT /*tag= a
FT FT /product= "SCA2"
FT FT /partial
FT FT /note= "the start and stop codons are not indicated"
XX
FN WO200289567-A2.
XX
PD 14-NOV-2002.
XX
PF 07-MAY-2002; 2002WO-US014698.
XX
PR 07-MAY-2001; 2001US-0289231P.
XX
PA (CEDA-) CEDARS SINAI MEDICAL CENT.
XX
PI Pulset SM;
XX
DR WPI; 2003-103467/09.
DR P-PSDB; ABB82696.
XX
XX New mutant mouse, useful as a disease model for identifying agents for
PT treating obesity or memory impairment, comprises a disrupted
PT spinocerebellar ataxia 2 gene.
XX
PS Disclosure; Page 52-54; 76pp; English.
XX
CC The invention relates to a mutant non-human mammal having a disrupted
CC spinocerebellar ataxia 2 (SCA2) gene. The mutant animal is useful as a
CC disease model for identifying agents or compounds that may treat obesity

CC or memory impairment. The stem cell is used for generating the mutant
CC animal. The present sequence represents the partial nucleotide sequence
XX (GenBank Accession No. U70670) encoding a partial mouse SCA2 polypeptide
SQ Sequence 1257 BP; 312 A; 330 C; 358 G; 257 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.8e-48 Length: 1257
Score: 1876.00 Matches: 365
Percent Similarity: 89.77% Conservative: 21
Best Local Similarity: 84.88% Mismatches: 18
Query Match: 26.95% Indels: 26
DB: 10 Gaps: 5

US-10-802-228-2 (1-1312) x ABV75289 (1-1257)

QY 157 GlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGln 176
Db 8 GGGCGCTCACCATGTCTGCTGAAGCCGACGCCGAG----- 43
QY 177 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
Db 44 -----CGCCCGCGCGCCCACTGGCGCG 67
QY 197 LysProGlyGlySerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSer 216
Db 68 AAGCCCGCGCGC--GGCCTCTCTCGTGGCGCGCGCGCGCGCGC---TCGGCGCGCG 121
QY 217 SerValSerSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSer----- 234
Db 122 GTGACCTCGGCTTCGGTGGTCCCGCGCGCGCGCGCGCGCGTCTCTCTCGCGCGCGC 181
QY 235 ---GlyGlyGlyArgProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGln 253
Db 182 CGGGCGCGCGCGCTCCCGGCTCGGCAGAGGTGGAAACAGTACCAAGGACTGCCTCAG 241
QY 254 SerThrIleSerPheAspGlyIleTyrAlaAlaMetArgMetValHisIleLeuThrSer 273
Db 242 CCTACGATTTCTTTTGATGGAATCTATGCAAAACGTGAGGATGTTTCTATATACCTTACGTCA 301
QY 274 ValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGluValPhe 293
Db 302 GTTGTGGATGAAATGTGAAGTACAAAGTGAACCAAGGAGCATATATGAAGGAGTCTTT 361
QY 294 LysThrTyrSerProLysCysAspLeuValLeuAspAlaAlaHisGlyLysSerThrGlu 313
Db 362 AAACATACAGTCTTAAGTGTGACTTGGTACTTGTCTGCACATGAGAAAAGTACAGAA 421
QY 314 SerSerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp 333
Db 422 TCCAGTTTCGGGGCCAAAACGTGAAGAAATAATGGAGAGTGTTTTGTTCAAATGCTCAGAC 481
QY 334 PheValValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThr 353
Db 482 TTCGTGTGGTACAGTTTAAAGATACAGACTCCAGTTATGACGAGAGAGTGTCTTTACT 541
QY 354 AspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTirp 373
Db 542 GACTCTGCTCAGCGCAAGGTGAATGGTGAGCACAAGGAGAGAGGACCTGGAGCCCTGG 601
QY 374 AspAlaGlyGluLeuThrAlaAsnGluGluLeuAlaLeuGluAsnAspValSerAsn 393
Db 602 GATGACGGGAGCTCACGGCCAGCGAGAGTGGAG---CTGGAGAATGATGTGTCTAAT 658
QY 394 GlyTirpAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThr 413
Db 659 GGATGGGACCCCAATGATGTCATATATGAAGAGAAATATGTTGTGTGTCTTCCACA 718
QY 414 TyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPhe 433
Db 719 TATGATAGCAGTTTATCTTCATATACGGTCTCTTAGAAGGGGACAACTCAGAGAAATTT 778
QY 434 LeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyr 453


```
QY 948 aGlyHisGlnProThrProValThr-----GlnProValCysPheAlaProA 966
Db 2397 -----CGGTCTGACAGCAGCCAGAGTGGCTATACAGCCCC 2436
QY 966 smMetMetTyrProValPro-----ValSerProGlyValGln-----ProLeuT 981
Db 2437 AGTACATCTCTTACATACCTCAGATCCATGGACAGCTGTGACGACCTCAGATGT 2496
QY 981 YrProileProMetThr---PrometProValAsnGlnAlaLysThrYrArgAlaValP 1000
Db 2497 ATCCATATCTGTATCCATTCAGTTCCTGGCAGCAGGCGCAAGTACCGGGGAGCAAAAG 2556
QY 1000 roAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAla-----MetMeth 1017
Db 2557 GCTCCCTCTCGCAGCAGCTCGGACCAACACCCAGCCAGCTCAGCCCGCCGATGATGC 2616
QY 1017 isProAlaSerAlaAlaGlyProPheAlaAlaThrProAlaLysThrGlnT 1037
Db 2617 AGCCCGCGGCTGTGCGCCGCTCTGTGTGCTGCCAGGCC---TATTCTTC---T 2670
QY 1037 YrValAlaTyrSerProGlnGlnPheProAsnGlnPro---LeuValGlnHisValProH 1056
Db 2671 ACATCCCTTACAACTCTAGCAGTTCAGGAGCCAGCCATGATGACGACCCATGGCCC 2730
QY 1056 isTyrGlnSerGlnHisProHisValTyrSerProValGlnGlyAsnAlaArgMetM 1076
Db 2731 ACTACCCCTCACAG-----CGGTGTTTGGCCCATGCTTTCAGACCAACCCAGCATGC 2784
QY 1076 etAlaProThrHisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyA 1096
Db 2785 TGACGTGGGAGCCATCCAG-----GCCATGTGTATCTCTACCTCAATACCTTT 2841
QY 1096 laHisGluGln-----ThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysG 1114
Db 2842 CTGACAGAGAGCTTACCCCAAGCCCTTATGCC----- 2876
QY 1114 luThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaH 1134
Db 2877 -----ACTGTTCCAGTCTTACCCAC 2898
QY 1134 isProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThrProThrG 1154
Db 2899 ACCATGCCACAGCTCCATGCCAC-----CAGCGCAGCGCGCTTACCCAGCCCTACTG 2952
QY 1154 yGlnGln---GlnSerGlnHisGlyGlySerHisProAlaProSerProValGlnHis- 1172
Db 2953 GAAGCAGCGCGAGTCCAGCATGCG-----GCCCCAGTCTGTCTCCAGGT-G 2999
QY 1172 ----- 1172
Db 3000 CTTGCCATGGGGGTGCTGAGTGTCTCTGTGAGGAATGGTGGCCAGAGAAGGATA 3059
QY 1172 ----- 1172
Db 3060 GAGCTAGGGGTCAATTTCTGAGTGGCGAGAGCTGGGGGCCAGCAGTGTGTGGCTGTGTG 3119
QY 1173 -----HisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerP 1186
Db 3120 GCACTCAACCTTCCCTCCCGCAGCAGCATCAGGCGGGGAGGCCCCACACTTGGGCGAGTG 3179
QY 1186 roGlnGlnGlnSerAlaIleTyrHisAlaGly---LeuAlaProThrProProSerMetT 1205
Db 3180 GACAGCCACAGCAAGATCTGTATCCACCCAGGGGCGCTGACAGGCGACCGCCCTCTCTGC 3239
QY 1205 hrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValP 1225
Db 3240 CACCGGAGCTTCTGCCAGCTGTCCTCAGAGAGCTTCCCTCC---CAGCCAGCGCGCTGTGT 3296
QY 1225 heThrIleHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisV 1245
Db 3297 ATGCCATCCACACAGCAGCTGCCCGCAGCGCTTCCACCAAC-----ATGCCCATG 3347
QY 1245 alProGlnAlaHisValGlnSerGlyMetVal-----ProSerHisProThrAlaH 1262
Db 3348 TTACCAGCCCATGTCCAACTGGAATCAGACAGAGCCCCCGCCCTCACCTTGGGCTC 3407
QY 1262 isAlaProMetMetLeuMetThrThrGlnProPro-----GlyGlyProGlnA 1278
Db 3408 CCCACCCCGCCCGAGTGTGTGTGTCACACCCAGAGTATGGGGGGCCCCCAAG 3467
QY 1278 laAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProT 1298
Db 3468 GCGCGGTGCCAGAGTGGGTGCTGCACTCTGACCTTCCACACCTCACCTTACCCAT 3527
QY 1298 YrMetThrHisPro 1302
Db 3528 ACATCGGACACCCC 3541
RESULT 14
ABZ72015
ID ABZ72015 standard; cDNA; 4674 BP.
XX AC ABZ72015;
XX DT 01-APR-2003 (first entry)
XX DE Human ataxin-2-like protein A2LP (A2LG) cDNA GenBank AF034373.
XX KW Human; cancer; stomach cancer; cytostatic; gene; ss.
XX OS Homo sapiens.
XX PN WO200283899-A1.
XX PD 24-OCT-2002.
XX PF 28-MAR-2002; 2002WO-JP003038.
XX PR 10-APR-2001; 2001JP-00112039.
XX PR 21-SEP-2001; 2001JP-00290193.
XX PA (TAKA-) TAKARA BIO INC.
XX PI Yoshikawa Y, Okamoto S, Oura T, Mineno J, Asada K, Kato I;
XX PI Inoue H, Mori M;
XX WI WI; 2003-093022/08.
XX DR Measuring changes in expression of 264 cancer associated genes for
XX PT detection of stomach cancer and screening of potential anticancer agents.
XX PS Claim 2; Page; 266pp; Japanese.
XX CC The invention relates to a method for the detection of cancer in which a
XX CC change in the expression of 1 or more of 264 specified cancer associated
XX CC genes, ABZ71694-ABZ71957, or of sequences at least 80% homologous to them
XX CC in the specimen tissue as compared to normal tissue is observed. The
XX CC genes are used in detection, diagnosis and treatment of cancer,
XX CC especially of stomach cancer. The present sequence is that of a cancer
XX CC associated polynucleotide of the invention. Note: The present sequence
XX CC was not given in the printed specification but was isolated using the
XX CC GenBank accession number given in the DE line
SQ Sequence 4674 BP; 971 A; 1566 C; 1244 G; 893 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4,58e-46 Length: 4674
Score: 1816.50 Matches: 504
Percent Similarity: 48.22% Conservative: 174
Best Local Similarity: 35.85% Mismatches: 381
Query Match: 26.10% Indels: 350
DB: 8 Gaps: 52
US-10-802-228-2 (1-1312) x ABZ72015 (1-4674)
QY 19 ArgPheAlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSer 38
```


Db	6	CGCTCCTC	CACTTCT	AGATGG	ATGGATG	CGCGCG	CGGGAAG	AGGCGCT	CCTCAT	TTTCT	CCTAGATG	65						
Qy	39	GlyArgGly	Gly	Gly	Ala	Ala	Pro	Gly	Pro	Tyr	Pro	Ser	Ala	Ala	Pro	Pro	Pro	68
Db	66	GGATGG	CGCGCG	CGGAGAG	TGCTCT	CAC	TTTCT	CCAGCC	CGGAGAC	CCCTCT	CCCTT	125						
Qy	59	GlyProGly	Pro	Pro	Pro	Ser	Arg	Gln	Ser	-----	Ser	Pro	Pro	Ser	Ala	Ser	Asp	76
Db	126	CCGCTC	CGCGCG	CTTCT	CGCGCG	GGT	TTTCT	CTCT	CCACCC	-----	-----	-----	170					
Qy	77	PheGly	Ser	Asn	Gly	Asn	Gly	Gly	Ala	Phe	Arg	Pro	Gly	Ser	Arg	Arg	Leu	96
Db	171	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	182					
Qy	97	LeuGly	Gly	Pro	Pro	Arg	Pro	Phe	Val	Val	Val	Leu	Leu	Pro	Leu	Ala	Ser	116
Db	183	-----	GGCT	CCCC	CGCGCG	-----	-----	-----	-----	-----	-----	-----	218					
Qy	117	Pro	Pro	Ala	Ala	Pro	Thr	Arg	Ala	Ser	Pro	Leu	Gly	Ala	Arg	Ala	Ser	136
Db	219	CCGATC	CCCTC	CGTTC	CGCGCG	-TCT	CCAGCG	GGGCCC	CAGCCCC	CGGCCCC	CTCT	CTC	277					
Qy	137	Gly	Val	Ser	Leu	Ala	Arg	Pro	Ala	Pro	Gly	Cys	Pro	Arg	Pro	Ala	Cys	156
Db	278	CTCC	CTTCT	CTA	-ATT	CCCTT	CCGAC	CGT-----	-----	-----	-----	-----	309					
Qy	157	Gly	Pro	Leu	Thr	Met	Ser	Leu	Leu	Pro	Gln	Gln	Gln	Gln	Gln	Gln	Gln	176
Db	310	-----	GGCAT	CAT	GTG	AAG	CCTC	AGCGCT	CAGCGCT	TAC	AAC	AGCC	TCC	CAG	CCCC	CAG	CAG	360
Qy	177	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Pro	Pro	Pro	Ala	Ala	Ala	-Asn	Val	Arg	196
Db	361	CGC	CCCC	ACG	CAAT	GTC	CGT	GCG	CGT	CGG	CCCC	CGG	GGG	GCG	CAC	GCCT	TCC	420
Qy	196	Gly	Pro	Gly	Gly	-----	-----	-----	-----	-----	-----	Ser	Gly	Leu	204			
Db	421	CGCG	CTC	CGCG	GGCG	GTG	GCAC	CTCT	GCG	GCCT	CTCT	CCG	GGCG	CTCC	CAG	CGCG	CGCT	480
Qy	204	uAla	Ser	Pro	Ala	Ala	-----	-----	-----	-----	Pro	Ser	Pro	Ser	Ser	Ser	Val	219
Db	481	CCCC	TGCT	GGCG	CGCT	GTG	CGCG	TGCG	GGAG	CGGG	CGTCC	CGCG	GGG	AGCG	CCG	AGG	CA	540
Qy	219	rSer	Ser	Ala	Thr	Ala	Pro	Ser	Ser	Val	Ala	Ala	Ala	Thr	Ser	Gly	Gly	239
Db	541	TCTT	GGCG	CGCG	CGCG	CGCG	CGCA	CAAC	CAC	GAG	GAG	GGCG	GGG	GGG	GAG	CGG	CA	600
Qy	239	o-Gly	Leu	Gly	Arg	Gly	Arg	Asn	Ser	Asn	Leu	Gly	Leu	Pro	Gln	Ser	Thr	259
Db	601	TCGG	CAGCG	CGCG	GAG	CAG	AGC	ACAG	GAA	GGG	AC	CCCC	CAC	AGT	CAC	CTGT	G	657
Qy	259	spGly	Leu	Tyr	Ala	Asn	Met	Arg	Met	Val	His	Leu	Thr	Ser	Val	Val	Gly	279
Db	658	AAGG	CGCT	CTA	CA	CA	AA	TTT	CC	AGA	ATG	CTG	CA	TTT	CTT	AC	AGT	717
Qy	279	ysGlu	Val	Gln	Val	lys	Asn	Gly	Gly	Leu	Tyr	Glu	Gly	Val	Phe	lys	Thr	299
Db	718	GTG	ATG	TAA	AGG	TG	AA	AT	TGG	TAC	CACT	TAT	GAG	GGT	AT	CTT	CA	777
Qy	299	ysCys	Asp	Leu	Val	Leu	Asp	Ala	Ala	His	Glu	Leu	Ser	Thr	Gln	Ser	Ser	319
Db	778	AGT	TTG	AAC	TAG	CGT	GTG	TGC	AC	CGG	AAA	AGC	ATC	TG	CAG	CAG	CGG	837
Qy	319	ysArg	Glu	Glu	Leu	Met	Gln	Ser	Leu	Phe	Leu	Cys	Ser	Asp	Phe	Val	Val	339
Db	838	GT	CGG	AGG	AG	CAT	TG	TG	GAC	CA	ATG	GGT	TTT					

QY 727 -----SerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGluA 745
 DB 1777 TGGGCTCTCCTCCAGCTTCTCCAAAGATCTCTCCGCCCCCACA-----1823
 QY 745 laLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnI 765
 DB 1824 -----GATGTAAAGAAGACTCTTACCNAAGNAACCTGGGAACTC 1863
 QY 765 leLysProAsnGlnThrSerProSerPheSerLysAlaGluAsnLysGlyLysSerProV 785
 DB 1864 TGGAGCCCCAGGAGCTG-----GCTCGGATAGCTGGGAAGTCCCTG 1905
 QY 785 alValSerGluHisArgLys-----GlnIleAspAspLeuLysLysPheLysAsnAspP 803
 DB 1906 GTCTTTCAGAAATGAACAGAACGATTCACACTGGAAAGACTGAAAGTTTGGGGCCCACT 1965
 QY 803 heArgLeuGlnProSerSerThrSerGlu--SerMetAspGlnLeuLeuAsnLysAsnA 822
 DB 1966 TTAAGCTTCAGCCCACTAGTCTCCCTGAGAACAGCTGGATCCTTTCCTCCCCGG-----2021
 QY 822 rgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysAspS 842
 DB 2022 -----ATCTTAAAGGAG-----GAGCCCAAGGAAAGGAGA 2052
 QY 842 erPheIleGluAsn-----SerSerSerAsnCyseT 852
 DB 2053 AAGAGCTTATGATGCTGTCTGTTGACTTCAGAGCCCATGGGGTCTCCCGTCTCTCCAAAGACAG 2112
 QY 852 hrSerGlySerLysProAsnSerProSerLysSerProSerLysLeuSerAsnThrG 872
 DB 2113 AGTCCGTATCGGATAGGAGGACAAACACCTCTGGCACCATCA-----GGAGGCACCTG 2166
 QY 872 luHisLysArgGlyProGluValThrSerGlnGlyVal-----GlnThrSerSerProA 890
 DB 2167 AG-----GGGCCAGAGCAGCCCCCACCACCTTGTCCAGCCCAAACTGGCAGCCCC 2217
 QY 890 laCysLysGlnGluLysAspAspLysGluGluLysLysAspAlaLaGluGlnValArgL 910
 DB 2218 CGGTGGGCTCATCAAGGAGGAGAGACAAAGATGAGGGCCCTGTGCTGAACAAGTAAAGA 2277
 QY 910 ysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSer-----PheSerGlnP 928
 DB 2278 AATCAACGTTGAACCCCTAATGCTTAAGGAGTTCAATCTCAAAAGCCTCTGCTGTCTGTGA 2337
 QY 928 roLysProSerThrProThrSerProArgProGlnAlaGlnProSerProSerMetV 948
 DB 2338 ATAAATCCACCACTACCCCACTTCTCCGGGGCCCGGACTCATTCAACTCCCTCCATC- 2396
 QY 948 alGlyHisGlnGlnProThrProValThr-----GlnProValCysPheAlaProA 966
 DB 2397 -----CCGGTGTGACAGCAGCCAGAGTGGGCTATACAGCCCCC 2436
 QY 966 snMetMetTyrProValPro-----valSerProGlyValGln-----ProLeuT 981
 DB 2437 AGTACATCTCTCATACACTCTCAGATCCATCGGACCAAGCTGTGAGGCACCTCAGATGT 2496
 QY 981 yrProIleProMetThr--ProMetProValAsnGlnAlaLysThrTyrArgAlaValP 1000
 DB 2497 ATCCATATCTGTATCAATTCATGCTGGGAGCAGGCAAGTACCGGGGAGCAAAAG 2556
 QY 1000 roAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAla-----MetMetH 1017
 DB 2557 GCTCCCTTCTCCGAGCGCTCGGACCAACACCAAGCAGCCTCAGCCCCCGCGATGATGC 2616
 QY 1017 isProAlaSerAlaAlaGlyProProIleAlaThrProProAlaLysThrSerThrGlnT 1037
 DB 2617 AGCGCGCGCGGTGTGTGGCCCGCTCTGTGGTGTGCTGACAGCC-----TATTCTTCC--T 2670
 QY 1037 yrValAlaLysSerProGlnGlnPheAsnGlnPro--LeuValGlnHisValProH 1056
 DB 2671 ACATCCCTCAACCCCTCAGAGTTCACAGCCAGCCAGCCATGATGATGACCCCATGGGCC 2730

QY 1056 isTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetM 1076
 DB 2731 ACTACCCCTCAG-----CCGGTGTGTGCCCCCATGCTTCAGAGCAACCCACCGCATGC 2784
 QY 1076 etAlaProProThrHisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyA 1096
 DB 2785 TGACGTTCGGGAGCATCCCCAG--GCCATCGTGTCTCTACCTCTCAATCACTCCCTT 2841
 QY 1096 laHisGluGln-----ThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysG 1114
 DB 2842 CTGAGAGCAGCCTACCCCAAGCCCTTATGTC-----2876
 QY 1114 luThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaH 1134
 DB 2877 -----ACTGTTACCCAGTCTTACCAC 2898
 QY 1134 isProAsnAlaThrLeuHisProHisThrProHisProGlnProSerSerAlaThrProThrG 1154
 DB 2899 ACCATGCCACACAGCTCCATGCCAC-----CAGCCGAGCGCGCTACCAACGCTACTGT 2952
 QY 1154 lyGlnGln--GlnSerGlnHisGlyGlySerHisProAlaProSerProValGlnHis- 1172
 DB 2953 GAAGCCAGCCGAGTCCAGCATGCG-----GCCCCAGTCTGTCTCAGGT-G 2999
 QY 1172 -----1172
 DB 3000 CCTGCCATGGGGGTGTGAGTGTCTGTGTCAGGAATGGTGGCCAGAGAGGGATA 3059
 QY 1172 -----1172
 DB 3060 GAGCTAGGGGTCAATTTCTGAGTGGCGAGACTGGGGGCCAGCGAGTGTGCTGTGTGTG 3119
 QY 1173 -----HisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerP 1186
 DB 3120 GCACCTCAACCTTCCCTCCACAGCAGCATCAGCGGGGCGGCGCCACACTTGGGCAGTG 3179
 QY 1186 roGlnGlnSerAlaIleTyrHisAlaGly--LeuAlaProThrProProSerMetT 1205
 DB 3180 GACGCCACAGCAGAAATCTGTATCCACCACCGGGGCCCTGACAGGACGCGCCCTCTCTGC 3239
 QY 1205 hrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValP 1225
 DB 3240 CACCGGGACCTTCTGCCAGTCCCTCAGAGCAGCTTCCCC--CAGCCAGCGCTGTCT 3296
 QY 1225 heThrIleHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisV 1245
 DB 3297 ATGCCATCCACCACAGCAGCTGCCCCACCGCTTCCCAAC-----ATGGCCCATG 3347
 QY 1245 alProGlnAlaHisValGlnSerGlyMetVal-----ProSerHisProThrAlaH 1262
 DB 3348 TTACCCAGGCCCATGTCCAACTGGAATCAGAGCAGCCCCCGCCCTCACCCTGGGGCTC 3407
 QY 1262 isAlaProMetMetLeuMetThrThrGlnProPro-----GlyGlyProGlnA 1278
 DB 3408 CCCACCGGCCAGGTGATGTCTGTGCCACCCACCCAGAGTATGGGGGGCCCCCACAAG 3467
 QY 1278 laAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProT 1298
 DB 3468 GCGGTGTGCCAGAGTGGGGTGCCTGTCTCTCAGCTTCCACACCTTCACTCCATCCCAT 3527
 QY 1298 yrMetThrHisPro 1302
 DB 3528 ACATCGGACACCCC 3541
 RESULT 15
 ADH29078
 ID ADH29078 standard; DNA; 4674 BP.
 XX
 AC ADH29078;
 XX
 DT 11-MAR-2004 (first entry)
 XX Human chronic myelogenous leukaemia (CML) gene marker #346.
 DE

476 snSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnIleProProG 496
1312 AGGGCTCAGGGCGGAGAGCCCGAGCTTGGCATCCAGGAGGAGTATATCCCTCTGC 1371
496 lylGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetG 516
1372 CTCAACGAGTCCGGGAAGT---CCCCGGGAGGAGTTGATGCACGACTCTCGGGGG 1428
516 lylGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAspPheAsn 536
1429 GTGGCTGGCCCTAGCTCTTTGCCACCTCGTGGCCCTCACCATCTGGACAACAGCAGCC 1488
536 roAsnSerGlySerAspGlnArgValIleAsnGlyValProTrpProSerProCysP 556
1489 CTGGCCAGGTTCTGAGGCCCGGTGATCAATGGA----- 1523
556 roSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProA 576
1523 ----- 1523
576 rgAlaAlaThrProThrArgProProSerArgProProSerArgProProS 596
1523 ----- 1523
596 erHisProSerAlaHisGlySerProAlaProValSerThrMetProLysArgMetSerS 616
1523 ----- 1523
616 erGluGlyProProArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValS 636
1524 -----GGCCCTTCCCGCATCTCCCAAGGACAGCGGCTCTGAGA----- 1565
636 erAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHisAsnProProSerG 656
1566 -----GGTGCACAGACTCTCTCTCG-----CCAGTA 1593
656 luAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThrTrpSerValV 676
1594 ATAGGCTTCGGAGAACTCTGTGCCACCTCTCTCGAGTGGG----- 1637
676 alSerGlyValProArgLeuSerProLysThrHisArgProArgProArgGlnAsnS 696
1638 -----CGGATGTATCCCGGTTCTCCCAAGTCTGCTGCCCT----- 1676
696 erIleGlyAsnThrProSerGlyProValLeuAlaSer-----ProGlnAlaG 712
1677 -----GCCCCAATCTCAGCTTCTGTCCAGAGCTCCCATCGGCT 1716
712 lylIleProThrGluAlaValAlaMetProIleProAlaAla----- 726
1717 CGGCAGTGCCAACTCTTCAGCCTCCATCCCTGTGACCTCATCAGTCTCAGATCTCGAG 1776
727 -----SerProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGluA 745
1777 TGGGCTCCATTTCTCCAGCTTCTCCAAAGATCTCCCTGGGCCCCACCA----- 1823
745 lalysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnI 765
1824 -----GATGTAAAGAACTCTCTTACCAAGAACTCTGGGAGAACTC 1863
765 leLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerProV 785
1864 TGGAGCCCGAGAGCTG-----GCTCGGATAGCTGGGAAAGTCCCTG 1905
785 alValSerGluHisArgLys-----GlnIleAspAspLeuLysLysPheLysAsnAsp 803
1906 GTCTTCAGATGAACAGAAACGATTCCAACTGGGAAGACTGAGAAAGTTTGGGCCCCAGT 1965
803 heArgLeuGlnProSerSerThrSerGlu---SerMetAspGlnLeuLeuAsnLysAsnA 822
1966 TTAAGCTTCAGCCAGTAGCTCCCTGAGAACACGCTGGATCTTTCTCCCGCG----- 2021
822 rgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysAspS 842

2022 -----ATCTTAAAGAG-----GAGCCCAAGAAAGGAGA 2052
842 erPheIleGluAsn-----SerSerSerAsnCyE 852
2053 AAGAGTTGAGTGGTCTGTGACTTCAGAGCCCATGGGTCTCCCGTCTCTCTCCAGACAG 2112
852 hrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnThrG 872
2113 AGTCCGTATCGATAGGAGGACAAACACCCCTGGCACCATCA-----GGAGGCACTG 2166
872 lylHisLysArgGlyProGluValThrSerGlnGlyVal-----GlnThrSerSerProA 890
2167 AG-----GGCCAGAGAGCCCCCACCCTTGTCCAAGCCAACTGGCAGGCCCC 2217
890 laCysLysGlnGluLysAspLysGluGluLysAspAlaAlaGluGlnValArgL 910
2218 CGGTGGGCTCATCAAGGAGAGAACAAAGATGAGGGCCCTGTGCTGAACAAGTAAAGA 2277
910 ysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSer-----PheSerGlnP 928
2278 AATCAACGTGTAACCCCTAATGCTAAGGAGTTCAATCTCAAAAGCTCTGCTGTCTGTGA 2337
928 roLysProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetV 948
2338 ATAAATCCACAGTACCCCAACTTCTCCGGGCGCCGGACTCATTCACACTCCCTCCATC- 2396
948 alGlyHisGlnGlnProThrProValTyrThr-----GlnProValCysPheAlaProA 966
2397 -----CCGGTCTCACAGCAGCCAGAGTGGGTATACAGGCCCC 2436
966 snMetMetTyrProValPro-----ValSerProGlyValGln-----ProLeuT 981
2437 AGTACATCTCTACATACCTCAGATCCATGAGGAGCAGCTGTGCAGGCACCTCAGATGT 2496
981 yrProIleProMetThr---ProMetProValAsnGlnAlaLysThrTyrArgAlaValP 1000
2497 ATCATATCTGTATCCAAATTCAGTGCCTGGCAGCAGGCAAGTACCGGGAGCAAAAG 2556
1000 roAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAla-----MetMetH 1017
2557 GCTCCCTTCTCCGAGAGCTCGGACCAACACAGCAGCCTCAGCCCCCGCGATGATGC 2616
1017 isProAlaSerAlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnT 1037
2617 AGCCCGCGGGCTGTGGCCGCTCTGTGGCTGCTGCCAGCCC---TATTCTTC---T 2670
1037 yrValAlaTyrSerProGlnGlnPheProAsnGlnPro---LeuValGlnHisValProH 1056
2671 ACATCCCTACAACTCAGCAGTTCAGGCGCAGCAGCCATGATGACGCCCATGGCCCC 2730
1056 isTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetM 1076
2731 ACTACCCCTCACAG-----CCGGTGTGGCCCCCATGTTTCAGAGCAACCCAGCATGC 2784
1076 etAlaProProThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyA 1096
2785 TGAGCTGGCGCAGCCATCCCCAG---GCCATCGTGTCTCTCTACCCCTCAATACCTT 2841
1096 laHisGluGln-----ThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysG 1114
2842 CTGAGAGCAGCTACCCCAAGCCCTTTATGCC----- 2876
1114 luThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaH 1134
2877 -----ACTGTTCCAGCTCTTACCAC 2898
1134 isProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThrProThrG 1154
2899 ACATGCCACACAGCTCCATGCCAC-----CAGCCGAGCGGCTACCCAGCCTACTG 2952
1154 lylGlnGln---GlnSerGlnHisGlyGlySerHisProAlaProSerProValGlnHis- 1172

```

Db      2953 GAAGCAGCGCGAGTCCCGAGCATGG-----GCCCCAGTCTGTCCAGGT-G 2999
Qy      1172 ----- 1172
Db      3000 CCTGCCATGGGGGTGCTGAGTGGTCTGTGTCAGGAATGGGTGGCCAGAGAAGGGATA 3059
Qy      1172 ----- 1172
Db      3060 GAGTAGGGGTCAATTCTGAGTGGCGAGGACTGGGGGCCAGCGAGTGTGCGCTGTGTG 3119
Qy      1173 -----HisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerP 1186
Db      3120 GCACCTCAACCTTCCCTCCCGCAGCAGCATCAGCGGGGGCGGCCCCACACTTGGGCAGTG 3179
Qy      1186 roGlnGlnGlnSerAlaIleTyrHisAlaGly---LeuAlaProThrProProSerMetT 1205
Db      3180 GACAGCCACACAGAGAATCTGTACCCAGCGGGGCCCTGACAGGCACGCGCCCTCTCTGC 3239
Qy      1205 hrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValP 1225
Db      3240 CACCGGGACCTTCTGCGCCAGTCCCTTCAGAGCAGCTTCCCC---CAGCCAGCGCGTGTGT 3296
Qy      1225 heThrIleHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisV 1245
Db      3297 ATGCCATCCACCACAGCAGCTGCCCGCGGCTTACCACAC-----ATGGGCCCATG 3347
Qy      1245 alProGlnAlaHisValGlnSerGlyMetVal-----ProSerHisProThrAlaH 1262
Db      3348 TTACCCAGGCCCATGTCCAACTGGAATCACAGCAGCCCCCGCCCTCACCTTGGGGCTC 3407
Qy      1262 isAlaProMetMetLeuMetThrThrGlnProPro-----GlyGlyProGlnA 1278
Db      3408 CCCACCGCCCGCCAGTGATGTGTGTCACACCCACCCAGAGTCATGGGGGGCCCCCACAAG 3467
Qy      1278 laAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProT 1298
Db      3468 GCGCGGTGCCCCAGAGTGGGGTGGCTGCACTCTCAGCTTCCACACCCCTACCCCTACCCAT 3527
Qy      1298 yrMetThrHisPro 1302
Db      3528 ACATCGGACACCCCC 3541

```

Search completed: December 23, 2005, 06:53:42
Job time : 1770 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 23, 2005, 05:15:22 ; Search time 9990 Seconds
(without alignments)
6144.610 Million cell updates/sec

Title: US-10-802-228-2
Perfect score: 6961
Sequence: 1 MRSAAAAPSPAVATESRRF.....TAHPPTMTSPVQAHHQOOL 1312

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlp
-O=/cgn2 1/USPTO.spool_p/US10802228/runat 21122005 110225 9713/app query.fasta_1.1479
-DB-EST -QPM-fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10802228 @CGN 1 1 8649 @runat 21122005 110225 9713 -NCPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2106.5	30.3	1528	4 CR619107	CR619107 full-length
2	1644.5	23.6	1060	1 AL554896	AL554896 AL554896
3	1589	22.8	977	7 CN646951	CN646951 ILLUMIGEN
4	1556	22.4	912	5 BQ883680	BQ883680 AGENCOURT
5	1530	22.0	938	5 BX375946	BX375946 BX375946
6	1525	21.9	876	3 BP159935	BP159935 BP159935
7	1508.5	21.7	1000	5 BX402110	BX402110 BX402110

8	1468	21.1	1100	3	BM455214	BM455214 AGENCOURT
9	1467.5	21.1	925	8	DR838409	DR838409 JG1 CABCT
10	1461.5	21.0	877	5	BQ428363	BQ428363 AGENCOURT
11	1448.5	20.8	859	5	BQ716397	BQ716397 AGENCOURT
12	1442.5	20.7	889	6	CA489164	CA489164 AGENCOURT
13	1430.5	20.6	917	8	DR867406	DR867406 JG1 CABG9
14	1410	20.3	830	6	CD616048	CD616048 56069370H
15	1403	20.2	865	8	CX568338	CX568338 UI-M-IBO-
16	1393.5	20.0	863	6	CD616049	CD616049 56069370J
17	1392.5	20.0	983	3	BI647561	BI647561 603278140
18	1386.5	19.9	1002	3	BM472410	BM472410 AGENCOURT
19	1383	19.9	846	1	AU124606	AU124606 AU124606
20	1366	19.6	871	5	BX411317	BX411317 BX411317
21	1364.5	19.6	845	1	AU118570	AU118570 AU118570
22	1353.5	19.4	813	6	CF739745	CF739745 UI-M-HDO-
23	1353	19.4	848	8	CX956610	CX956610 JG1 CAA08
24	1351	19.4	831	3	BM948018	BM948018 UI-M-EGOP
25	1344	19.3	764	3	BP149122	BP149122 BP149122
26	1340.5	19.3	924	5	BU513989	BU513989 AGENCOURT
27	1329	19.1	800	3	BP157220	BP157220 BP157220
28	1327.5	19.1	865	7	CK310762	CK310762 S802006B2
29	1316	18.9	818	3	BI647389	BI647389 603279646
30	1308	18.8	750	1	AU142284	AU142284 AU142284
31	1303.5	18.7	806	7	CO573665	CO573665 AGENCOURT
32	1299.5	18.7	808	6	CF256498	CF256498 mdvnl40 e
33	1299.5	18.7	940	5	BU159977	BU159977 AGENCOURT
34	1299	18.7	749	6	CD616053	CD616053 56069494J
35	1297	18.6	753	7	CN786484	CN786484 4120484 B
36	1291.5	18.6	807	6	CB244158	CB244158 UI-M-PYO-
37	1289	18.5	779	1	AU121620	AU121620 AU121620
38	1278	18.4	969	6	CF237560	CF237560 AGENCOURT
39	1269	18.2	791	6	CF537566	CF537566 UI-M-G10-
40	1264.5	18.2	970	2	BG254942	BG254942 602369395
41	1261	18.1	860	6	CD656309	CD656309 AGENCOURT
42	1256	18.0	745	8	CX218145	CX218145 MNS32508
43	1249	17.9	717	5	BX501120	BX501120 DKFP2779A
44	1248	17.9	684	6	CD616052	CD616052 56069494H
45	1246	17.9	697	6	CD616051	CD616051 56069478J

ALIGNMENTS

RESULT 1	CR619107	1528 bp	mRNA	linear	HTC 21-JUL-2004
LOCUS	full-length cDNA clone CS0D1087Y117 of Placenta Cot 25-normalized				
DEFINITION	of Homo sapiens (human)				
ACCESSION	CR619107				
VERSION	CR619107.1 GI:50499914				
KEYWORDS	HTC; CNSLT_CDNA				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1. (bases 1 to 1528)				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished				
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1528) Genoscope.				
REFERENCE	Direct Submission				
AUTHORS	Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :				
TITLE	BP 191 91006 EYRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
JOURNAL	- Web : www.genoscope.cns.fr)				
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.				


```

FEATURES
  source      Location/Qualifiers
1..1528
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="CSODI087Y117"
  /tissue_type="Placenta Cot 25-normalized"
  /plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.:      3,85e-100      Length:      1528
Score:          2106.50        Matches:      395
Percent Similarity: 92.18%      Conservative: 6
Best Local Similarity: 90.80%    Mismatches:  19
Query Match:     30.26%        Indels:      15
DB:              4            Gaps:        3

US-10-802-228-2 (1-1312) x CR619107 (1-1528)
QY      861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
DB      1 CCCAGCAITTCCTCCCTCAATACTTAGTAACACGGAGCACAGAGGGGACCTGAGGTCACT 60
QY      881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
DB      61 TCCCAAGGGGTTACAGATTCAGCCCGCAGCATGTAAACAAGAGAAAGACGATAAGGAAGAG 120
QY      901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
DB      121 AAGAAGACGACGCTGAGCAAGTTAGGAATAACAATTGAATCCCAATGCCAAGAGAGTTTC 180
QY      921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
DB      181 AACCCACGTTCTTCTCAGCCAAAGCCTTCTACTACCCCACTTCACTCCGGCTCAA 240
QY      941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
DB      241 GCACAACCTAGCCCATCTATGGTGGGTATCAACAGCCAACTCCAGTTTATATCTCAGCCT 300
QY      961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
DB      301 GTTGTGTTTGGACCAAAATATGATGATCCAGTCCAGTCCAGTCCAGCCGCGGTGCAACTTTA 360
QY      981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAla----- 998
DB      361 TACCCAATACCTATGACGCCCATGCGCAGTGAATCAAGCCACAGACATATAGACGAGTAAA 420
QY      999 ValProAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisPro 1018
DB      421 GTACCAAAATATGCCCAACAGCGGCAAGACCAAGCATCATCAGATGCCATGATGCACCA 480
QY      1019 AlaSerAlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrVal 1038
DB      481 GCGTCAGCAGCGGGCCCAACCGATTCCAGCCACCCACAGCTTACTCCACCAATATGTT 540
QY      1039 AlaTyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGln 1058
DB      541 GCCTACAGTCTCTCAGAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAG 600
QY      1059 SerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaPro 1078
DB      601 TCTCAGCATCTCATGTCTATAGTCTGTAAATACAGGGTAAATGCTAGAATGATGCCACCA 660
QY      1079 ProThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGlu 1098
DB      661 CCAACACACGCCCGCCCTGGTTTATGATCTCTTCAGCAACTCAGTACGGGCTCATGAG 720
QY      1099 GlnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSer 1118
DB      721 CAGAGCATGCATGTATGCATGTCTCCAAATTTACATACACAGAGACACAGCCCTTCT 780
QY      1119 PheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThr 1138

```

```

DB      781 TTCTACTTTGCCATTTCCACGGGTCTCTTCTCAGCAGTATGCACACCTTAAGCTACC 840
QY      1139 LeuHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSer 1158
DB      841 CTGACCCACATATCTCCACACCTCAGCCTTACGCTTACCTCCCACTGGACACAGCAAAGC 900
QY      1159 GlnHisGlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAla 1178
DB      901 CAACATGTTGAAGTCTATCTCTGCACCCAGTCTCTGTTCAGCACCATCAGCACCAGCGCCGC 960
QY      1179 GlnAlaLeuHisLeuAlaSerProGlnGlnSerAlaLalleTyrHisAlaGlyLeuAla 1198
DB      961 CAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTCCAGCAATTTACCAACGCGGGCTTCG 1020
QY      1199 ProThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPhePro 1218
DB      1021 CCAACTCCACCTCCATGACACCTGCTTCCAAACAGCAGTCCGACAGATAGTTTCCCA 1080
QY      1219 AlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsn 1238
DB      1081 GCAGCACAAACAGACTGTCTTTACGATCCATCTCTTCAGCTTCAGCGCGGTATACCAAC 1140
QY      1239 ProProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHis 1258
DB      1141 CCACCCCAACATGGGCCACGTACCTCAG-----TACAAG 1173
QY      1259 ProThrAlaHisAlaProMetMetLeuMetThrThrGlnProGlyGlyProGlnAla 1278
DB      1174 CCCACCACCAACAGAGTTGTAAGCTGCTCCCTGGAGGAACCAAGGC----- 1221
QY      1279 AlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThr 1293
DB      1222 CAAATTCCTCTCTCTCTCTACTGCTTCTACCACTGGAAGCACA 1266

RESULT 2
AL554896      1060 bp      mRNA      linear      EST 30-MAR-2004
LOCUS      AL554896 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION      clone CSODI087Y117 5-PRIME, mRNA sequence.
ACCESSION      AL554896
VERSION      AL554896.3 GI:45859645
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE      1 (bases 1 to 1060)
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      On Feb 15, 2001 this sequence version replaced gi:31276706.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6451.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODI087AE09QPI&c=6451.r.
Location/Qualifiers
1..1060
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="CSODI087Y117"
  /tissue_type="PLACENTA COT 25-NORMALIZED"
  /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

FEATURES
  source

```

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:

Pred. No.: 3,47e-76 Length: 1060
 Score: 1644.50 Matches: 319
 Percent Similarity: 90.58% Conservative: 8
 Best Local Similarity: 88.37% Mismatches: 21
 Query Match: 23.62% Indels: 15
 DB: 1 Gaps: 4

US-10-802-228-2 (1-1312) x AL554896 (1-1060)

QY	861	ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr	880
Db	1	CCGAGCATTTCCCTTCAATACCTAGTAAACAGGAGCAAGAGGGAGCTGAGTCACT	60
QY	881	SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGlnLysAspLysGluGlu	900
Db	61	TCCCAAGGGTTCCAGATTTCCAGCCCGCATGTATAACAGAGAAAGCATAGGAAGAG	120
QY	901	LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe	920
Db	121	AGAAAGAGCGCAGCTGAGCAAGTMAAGAAATCAACATTCGAATCCCAATGCAAGAGATTC	180
QY	921	AsnProArgSerPheSerGlnProLysProSerThrProThrSerProArgProGln	940
Db	181	AACCCAGTTCTTCTCTCAGCAAGCCCTTCTACCCCACTTACCTCGGCTCAA	240
QY	941	AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro	960
Db	241	GCACACCTAGCCCATCTATGTTGGTTCATCAACAGCCCACTCCAGTTTACTACGCT	300
QY	961	ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu	980
Db	301	GTCTGTGTTGCACCAATATGATGATATCCAGTCCAGTCCAGTCCAGTCCAGTCCAG	360
QY	981	TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAla	998
Db	361	TACCCCAATACCTATGAGCCCATGCGCATGTAATCAAGCCCAAGACATATAGAGCGTAA	420
QY	999	ValProAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisPro	1018
Db	421	GTACCAATATGCCCCACAGCGCGCAGACAGCATCATCAGAGTGCCATGATGCCCA	480
QY	1019	AlaSerAlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrVal	1038
Db	481	GCCTGACA-CCGGGCCCCAGATTGCGACCCACCCAGCTTACTCCAGCAATATGTT	539
QY	1039	AlaTyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGln	1058
Db	540	GCCTACAGTCTCTCAGCAGTACCCCAATACGCCCCCTTTTCCAGCATGTGCCACATATCAG	599
QY	1059	SerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaPro	1078
Db	600	TCTCAGCATCTCTATGCTATGATCTCTGTAATACAGGGTAATGCTAGATGATGGCACCA	659
QY	1079	ProThrHisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGlu	1098
Db	660	CCAAACACACCCAGCGCTGTTTAGTATCTCTTCCAGCACTCAGTACGGGGCTCATGAG	719
QY	1099	GlnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSer	1118
Db	720	CAGACGATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	779
QY	1119	PheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThr	1138
Db	780	TTCTACTTTGCAATTTCCAGGGCTCTCTGCTCAGCAGTATGCGACCCCTTACGCTACC	839
QY	1139	LeuHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSer	1158

Db	840	CTGCACCCACATCTCCACACCTCAGCTTCCAGTACCTCCCTCCCTGGACAGCAARC-	898
QY	1159	GlnHisGlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAla	1178
Db	899	CAACATGTTGGAAGTCACTCTGCACCAAGTCTGTTTCCAGCAGCAGGCGGCC	958
QY	1179	GlnAlaLeuHisLeuAlaSerPro-----GlnGlnGlnSerAlaIleTyrHisAlaGly	1196
Db	959	-----AGCTCTCCATYKCCAGTCAMAGCAGCTTACGATTTTACAMGGGG	1003
QY	1197	LeuAlaProThrProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSer	1216
Db	1004	GSTGCG-----CACTCCTCAGAACTCTCTCMAAAVAGATGCAAAAATAT	1051
QY	1217	Phe 1217	
Db	1052	TTC 1054	
RESULT 3			
CN646951			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
COMMENT			

FEATURES

source

1. .977
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /strain="Indian"
 /db_xref="taxon:9544"
 /clone="IBIUM:8082"
 /sex="female"
 /dev_stage="adult"
 /lab_host="E. coli SOLR"
 /clone_lib="Katze_MBR"
 /notes="Organ: brain; Vector: Uni-ZAP XR; Site: 1; EcoR I; Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis Kit (Catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (Catalog #200450)"

ORIGIN

Alignment Scores:

QY 909 gtysserThrLeuAsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProIly 929
 Db 663 GAAATCAACATGGAATCCCAATGCAAGAGAGTCAACCCACGTTCTCTTCAGCCAAA 722
 QY 929 sProSerThrProThrSerProArgProGlnAlaGlnProSerProSerMetValG1 949
 Db 723 GCCTTCTACTACCCCACTTCACCTCGGCTCAAGCAACACTAGCCCATCTATGGTGGG 782
 QY 949 yHisGlnGlnProThrProValTyrThrGlnProValCysPheAlaProAsnMetMetTy 969
 Db 783 TCATCAACAGCCCAAMTCCAGTTTATATCTCAGCTCTGTTTGTGCAACCAATATGATGA 842
 QY 969 rProValProValSerProGlyValGlnProLeuTyrProIle-ProMetThrProMetP 989
 Db 843 TWCAGTCCAGTGAGCCCGGCGTCAACCTTTATACCCMATTTACCTATGACGCCCATGC 902
 QY 989 roValAsnGlnAlaLysThrTyrArgAla 998
 Db 903 CAGTGAATCAAGCCAGACATATAGCA 931

RESULT 6
 BP159935 876 bp mRNA linear EST 30-DEC-2003
 LOCUS BX402110
 DEFINITION BX402110 Homo sapiens enriched swine cDNA library, adult thymus Sus
 scrofa cDNA clone THY010091A06 5', mRNA sequence.
 ACCESSION BP159935
 VERSION BP159935.1 GI:40409408
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa

REFERENCE
 AUTHORS Uenishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H., Okumura, N., Hamasima, N. and Awata, T.
 TITLE PEDE (Fig EST Data Explorer): Construction of a database for ESTs derived from porcine full-length cDNA libraries
 JOURNAL Nucleic Acids Res. 32 (1), D484-D488 (2004)
 PUBMED 14681463

COMMENT
 Contact: Hirohide Uenishi
 Animal Genome Laboratory, Genome Research Department
 National Institute of Agrobiological Sciences
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
 Tel: +81-29-838-8627
 Fax: +81-29-838-8627
 Email: huenishi@affrc.go.jp
 EST project with full-length enriched cDNA libraries carried out in Animal Genome Research Program (Japan) by National Institute of Agrobiological Sciences and STAFF-Institute
 Single pass sequencing of clones derived from oligo-capped cDNA library

Vector sequences were eliminated by RepeatMasker version 2002/07/13 and crossmatch version 0.990319
 Low quality bases were trimmed based on the quality values.

FEATURES

source
 Location/Qualifiers
 1..876
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="THY010091A06"
 /issue_type="thymus"
 /dev_stage="adult"
 /clone_lib="full-length enriched swine cDNA library, adult thymus"

ORIGIN

Alignment Scores:
 Pred. No.: 4,79e-70 Length: 876
 Score: 1525.00 Matches: 283
 Percent Similarity: 92.26% Conservative: 3
 Best Local Similarity: 91.23% Mismatches: 4
 Query Match: 21.91% Indels: 20

DB: 3 Gaps: 2
 US-10-802-228-2 (1-1312) x BP159935 (1-876)
 QY 991 AsnGlnAlaLysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGlnAsp 1008
 Db 1 AATCAAGCCAAGACATATAGCAGAGGTAAAGTACCAATATATACCCCAACACGCGAAGAC 60
 QY 1009 GlnHisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAla 1028
 Db 61 CAGCATCATCAGAGCACCATGATGCACCCAGCCTCAGCAGAGGTCCACCAATTGTTGCC 120
 QY 1029 ThrProProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsnGln 1048
 Db 121 ACCCCAGCAGCTTATCCACACAATATGTCCTATATAGTCTCAGCAGATTTCCAAATCAA 180
 QY 1049 ProLeuValGlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerProVal 1068
 Db 181 CCTCTTGTTCAGCATGTGCCACATTCATGCTCAGCATCTCATGTCTATAGTCCTGTGA 240
 QY 1069 IleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSer 1088
 Db 241 ATACAGGGTAATGCTAGAAATGATGGCACCAACACATGCCAGCCTGGTTTATGATCT 300
 QY 1089 SerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysProLys 1108
 Db 301 TCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACACATGCGATGTAT----- 348
 QY 1109 LeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeu 1128
 Db 349 -----GTTTCCACTGGCTCCCTT 366
 QY 1129 AlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnPro 1148
 Db 367 GCTCAACAGTATGCGACCCCTAATGCTACCTGCACCCACATATCTCCACATCTCCTCAGCCT 426
 QY 1149 SerAlaThrProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaProSer 1168
 Db 427 TCAGTACTCTCCACTGGACAGCAGCAGCAACCGCATGGTGAAGCCATCTCTGCACCCAGT 486
 QY 1169 ProValGlnHisHisGlnHisGlnAlaGlnAlaLeuHisLeuAlaSerProGlnGln 1188
 Db 487 CCTTTTCAGACACCATCAGCACCGCTGCCAGGCTCTCCATCTGGCCAGTCCACACAG 546
 QY 1189 GlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProAlaSer 1208
 Db 547 CAGTCAGCCATTACCATGCGAGGCTCGTCCACACCCACCTTCATGACACCTGCCTCC 606
 QY 1209 AsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIleHis 1228
 Db 607 AATACGCAGTCGCCACAGATAGCTTCCCAACAGCACAACAGACTGTCTTCACGATCCAT 666
 QY 1229 ProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisValProGlnAla 1248
 Db 667 CTTTCTCATGFTCAGCCGGCATACCAATCCACCCACATGGCCCATGTACCTCAGCCT 726
 QY 1249 HisValGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeuMet 1268
 Db 727 CATGTACAGTCAGGAATGGTTCTTCTCATCTCAACTGCCCATGGCCCAATGATGCTAATG 786
 QY 1269 ThrThrGlnProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnProIle 1288
 Db 787 ACGACACAGCCACCCGGCGTCCCGCAGCGCCCTCGCTCANAGTGCACTACAGCCCAT 846
 QY 1289 ProValSerThrThrAlaHisPheProTyr 1298
 Db 847 CCAGTCTCGACACAGCGCATTTCCCTTAT 876

RESULT 7
 LOCUS BX402110
 DEFINITION BX402110 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DC015YB14 5-PRIME, mRNA sequence.
 1000 bp mRNA linear EST 01-MAY-2004


```

source      1..1100
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500163"
/tissue_type="lymphoma, cell_line"
/lab_host="PH108 (phage-resistant)"
/clone_lib="NIH_MGC_85"
/notes="Organ: lymph; Vector: pCMV-SpORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.:      6,21e-67      Length:      1100
Score:          1468.00      Matches:      301
Percent Similarity: 95.31%      Conservative: 4
Best Local Similarity: 94.06%      Mismatches: 8
Query Match:      21.09%      Indels:      7
DB:              3          Gaps:      2

US-10-802-228-2 (1-1312) x BM455214 (1-1100)

QY 74 SerAspCysPheGlySerAsnGlyAsnGlyGlyAlaPheArgProGlySerArgArg 93
Db 3 TCAGACTGTTTGGTAGCAACGCAACGCGCGCGGCTTCGGCGCGGCTCCCGCGG 62

QY 94 LeuLeuGlyLeuGlyGlyProProArgProPheValValValLeuLeuProLeuAlaSer 113
Db CTTCTTGTTGCTCGGCGGCGCTCCCGCGCGCTTCGTGCTCTCTCTCTCTCTCTCTCT 122

QY 114 ProGlyAlaProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAlaSerPro 133
Db CCGGCGCGCCCTCCCGCGCGCGCAACCGCGGCTCCCGCTCGGCGCGCGCGGCTCCCG 182

QY 134 ProArgSerGlyValSerLeuAlaArgProAlaProGlyCysProArgProAlaCysGlu 153
Db CCGCGTTCCGGGCTCTCTTGGCGCGCGCGGCTCCCGGCTGTCCTCCCGCGCGGTGCGAG 242

QY 154 ProValTyGlyProLeuThrMetSerLeuLysPro---GlnGlnGlnGlnGlnGln 172
Db CCGGTTGTATGGGCGCCCTCACCATGTGCTGTGAAGCCCGACGACGACGACGACGACG 302

QY 173 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 192
Db CAAACAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 362

QY 193 AlaAsnValArgLysProGlyGlySerGlyLeuLeuAlaSerProAlaAlaAlaProSer 212
Db GCCAATGTCCGCAAGCCCGCGGCGCGGCTTCCTAGCGTCCCGCGCGCGCGGCTTCG 422

QY 213 ProSerSerSerValSerSerSerSerAlaThrAlaProSerSerValValAlaAla 232
Db CGTCTCTGCTCTCGGCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 482

QY 233 ThrSerGlyGlyGlyArgProGlyLeuGlyArgGlyValGlyAsnSerAsnLysGlyLeuPro 252
Db ACCTCGGCGGCGGAGGCGCGCGCTGGGCGAGAGTCTGAAACAGTAACAAAGGACTGGCT 542

QY 253 GlnSerThrIleSerPheAspGlyIleTyAlaAsnMetArgMetValHisIleLeuThr 272
Db CAGTCTACGATTCTTTTGGTGAATCTATGCAATATATGAGGATGTTTCATATCTTACA 602

QY 273 SerValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrcGlyVal 292
Db TCAGTTGTGGCTCCAAATGTGAAGTACAAGTGAAGAAATGGAGGTATATATGAAGGAGTT 662

QY 293 PheLysThrTySerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThr 312
Db TTTNAACTTACAGTCCGAGTGTGATTGGTACTTGTGATGCCGCACATGAGNAAAGTACA 722

```

```

QY 313 GluSerSerSerGlyProLysArgGluGluLeuMetGluSerIleLeuPheLysCysSer 332
Db GAATCCAGTTCCGGGCCGAAACGCTGAAGAAATAATGGAGAGTATTTTGTTCATGTTCA 782

QY 333 AspPheValValValGln-PheLysAspMetAspSerSerTyAlaLysArgAspAlaPh 352
Db GACTTTGNTGTGTACAGTTTAAAGATATGAGCTCCAGTTATGCAAAAAGAGATGCTTT 842

QY 352 eThrAspSerAlaIleSerAla-LysValAsnGly-GluHisLysGluLysAspLeu-GI 371
Db TACTGACTCTCTATCATGCTGTAAGTGAATGGCGGACCCCAAGAGAGAGACCTGGNA 902

QY 371 uProTrp-AspAlaGlyGluLeuThrAlaAsn---GluGluLeuGluAla 386
Db GGCCTGGGATGCGAGGGAACCTCCAGCCCAATTAGGGAACCTTGAAGCT 952

RESULT 9
DR838409 925 bp mRNA linear EST 28 JUL 2004
LOCUS JGI CAB7383.fwd NIH XGC tropFatl Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7803839 5', mRNA sequence.
ACCESSION DR838409
VERSION DR838409.1 GI:71457349
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 925)
Richardson,P., Lucas,S., Rokhsar,D., Dettter,J.C., Ng,D.C.,
Brokstein,P. and Lindquist,E.A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
Other ESTs: JGI CAB7383.rev
Contact: Lindquist,E.A., Richardson,P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Bruce Blumberg Laboratory, University of
California, Irvine
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
Clone Distribution: I.M.A.G.E. Consortium/LNL.
http://image.llnl.gov
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Place: CAB 0077 row: m column: 21
High quality sequence stop: 871.
Location/Qualifiers
1..925
/mol_type="mRNA"
/organism="Xenopus tropicalis"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7803839"
/dev_stage="Adult"
/tissue_type="Fat body"
/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
/clone_lib="NIH XGC tropFatl"
/notes="Vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The
library was prepared from 5 ug of poly A+ RNA by oligo dT
priming
(5'-ACTAGTCGCGCCCTAGGCTCGAGTTTTTTTTTTTTTTTTTTT-3') and
Stratascript reverse transcriptase. After ligation of
EcoRI adapters (5'-AATTCGGCAGGAG-3') followed by kinasing
adapters and by XhoI digestion, the cDNA was size selected
by chromatography on Sepharose CL-2B columns and fractions
containing cDNAs larger than 1000 bp were ligated into

```

FEATURES
source

EcoRI/XhoI-digested pCS107. Reference for library construction: Current Genomics 4, 635-644. Library constructed by Michelle Tabb and Bruce Blumberg (Dept of Developmental and Cell Biology, University of California, Irvine).

ORIGIN

Alignment Scores:
 Pred. No.: 5,24e-67 Length: 925
 Score: 1467.50 Matches: 272
 Percent Similarity: 86.20% Conservatives: 9
 Best Local Similarity: 83.44% Mismatches: 24
 Query Match: 21.08% Indels: 21
 DB: 8 Gaps: 3

US-10-802-228-2 (1-1312) x DR838409 (1-925)

QY	934	ProThrSerProArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnPro	953
DB	3	CCAACTCAGCTCGTCCCAAACTCAGCAAGTCCATCATGTTGTGGACATCAGCAGCCA	62
QY	954	ThrProValThrThrGlnProValCysPheAlaProAsnMetMetTyrProValProVal	973
DB	63	ACGCTGTGTATACCCAGCAGCATGTTTTCACCGCAATATGATGATCCAGTCTCTGTG	122
QY	974	SerProGlyValGlnProLeuTyrProLeuProMetThrProMetProValAlaGlnAla	993
DB	123	AGCCTGGAGTTCAGCCATATATCTCCATTCATGACTACATGAGTCCAGTGAACAGGCC	182
QY	994	LyethTyrArgAla-----ValProAsnMetProGlnGlnArgGlnAspGlnHis	1011
DB	183	AAGACTTACAGACAGCATGAAGTACCAATATATCCCGCAAGTCCAGCAGCCAC	242
QY	1012	GlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProLeuAlaAlaThrProPro	1031
DB	243	CAGAAATCATGATGATCATCAGTATCTGACGAGGAGCCACCAATAGTTGTACCCACCT	302
QY	1032	AlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsnGlnProLeuVal	1051
DB	303	GCCTATTCTGCCAGTATGCTGCTTACAGCCCTCAACAGTTTCTTAATCAACCACTTATG	362
QY	1052	GlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGly	1071
DB	363	CAGCATGTACAGCACTATCATGTCACAGCACCCCATGTATATATAGTCTGTGATACAGGG	422
QY	1072	AsnAlaArgMetMetAlaProThrHisAlaGlnProGlyLeuValSerSerAla	1091
DB	423	AACACAGAATGATGGCAGCCACCATCTCATGACAGCTGGCTTGTATCTCTCGCT	482
QY	1092	ThrGlnTyrGlyAlaHisGlnGlnThrHisAlaMetTyrAlaCysProLysLeuProTyr	1111
DB	483	GCACAGTATGCCACTCTCGAACAACCTCATACCATGTAT-----	521
QY	1112	AsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGln	1131
DB	522	-----GTTTCAACTAGCTCATTAGCAGCAG	548
QY	1132	TyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThr	1151
DB	549	TATGCTCACCACAAATGAGCCCTGACCCACCCATCTCATCTCATCTCAGCCTTACGAACT	608
QY	1152	ProThrGlyGlnGlnSerGlnHisGlyCysSerHisProAlaProSerProValGln	1171
DB	609	CTTACCGGCCACAGCAAGTCAACATGGGGAGTCAACCCAGTCCCAAGTCCAGTTTCAG	668
QY	1172	HisHisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerProGlnGlnSerAla	1191
DB	669	CACCATCAACATCAGGCATCACAAGCTCTGCATCTGGCTAACCAACAGCAGCAGTCTGCA	728
QY	1192	IleTyrHisAlaGlyLeuAlaProThrProSerMetThrProAlaSerAsnThrGln	1211
DB	729	ATTTTACCATGAGGACTTGCTCCAAACGCCCTCTCCCATGACACCAAGCTCCACGCA	788

QY	1212	SerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIleHisProSerHis	1231
DB	789	TCCCAAGAGAGAGCTTCCCA---ACTCAACAGAGCTGTGTATACCATTCACCATCACAT	845
QY	1232	ValGlnProAlaTyrThrAsnProProHisMetAlaHisValProGlnAlaHisValGln	1251
DB	846	GTCCAAGCAGCATATACCAATCCGCTCACATGGCCCATGTGCAACAGCAGCATGTACAG	905
QY	1252	SerGlyMetValProSer	1257
DB	906	TCTGGAATGGTTCCTTCT	923
RESULT 10			
BQ428363			
LOCUS			
DEFINITION BQ428363 877 bp mRNA linear EST 24-MAY-2002			
5', mRNA sequence.			
ACCESSION BQ428363			
VERSION BQ428363.1 GI:21167439			
KEYWORDS EST.			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE 1 (bases 1 to 877)			
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.			
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL Unpublished (1999)			
COMMENT Contact: Robert Strausberg, Ph.D.			
Email: cgaabp-r@mail.nih.gov			
Tissue Procurement: ATCC			
cDNA library Preparation: Life Technologies, Inc.			
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Agencourt Bioscience Corporation			
Clone distribution: MGC clone distribution information can be			
found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLAM13549 row: h column: 12			
High quality sequence stop: 652.			
FEATURES			
-source			
1..877			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:6175163"			
/tissue_type="leiomyosarcoma"			
/lab_host="DH10B (phage-resistant)"			
/clone_lib="NIH_MGC_71"			
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."			
ORIGIN			
Alignment Scores:			
Pred. No.: 1.01e-66 Length: 877			
Score: 1461.50 Matches: 279			
Percent Similarity: 91.56% Conservatives: 3			
Best Local Similarity: 90.58% Mismatches: 4			
Query Match: 21.00% Indels: 22			
DB: 5 Gaps: 2			
US-10-802-228-2 (1-1312) x BQ428363 (1-877)			
QY	880	ThrSerGlnGlyValGlnThrSerSerProAlaCysLysGlnGlyLysAspLysGlu	899
DB	1	ACTTCCCAAGGGTTGAGCTTCCAGCCAGCATGTAAACAAGAAAGACGATAAGAA	60
QY	900	GluLysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGlu	919
DB	61	GAGAGAAAGACGAGCTGAGCAAGTTAGGAATCAACATTGAATCCCATGCAAGAG	120
QY	920	PheAsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgPro	939

```
Db 121 TTCAACCCACGTTCTCTCTCAGCAAGCCCTTCTACTACCCCACTTCACTCGGCT 180
Qy 940 GlnAlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGln 959
Db 181 CAAGCACACCTAGCCCACTATGTGTGGTTCATCAACAGCCCACTCCAGTTTACTCTAG 240
Qy 960 ProValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnPro 979
Db 241 CTTGTTGTGTTTGCACAAATATGATGTATCCAGTCCCACTCCAGTCCAGCCGCGCAACT 300
Qy 980 LeuTyrProIleProMetThrProMetProValAsnGlnAlaValThrTyrArgAlaVal 999
Db 301 TTATACCCCAATACCTATGAGCCCATGCGGAGTCAAGTCAAGCCCAACATATAGAGCA 360
Qy 1000 ProAsnMetProGlnGlnArgGlnAspGlnHisGlnHisGlnSerAlaMetMetHisProAla 1019
Db 361 CCAATATGCCCAACAGCGGCAAGCAGCATCATCAGATGCCATGATGACCCAGCG 420
Qy 1020 SerAlaAlaGlyProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAla 1039
Db 421 TCAGCAGCGGGCCCAACGATGTCAGCCACCCACAGCTTACTCCAGCAATATGTTGCC 480
Qy 1040 TyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSer 1059
Db 481 TACAGTCCTCAGCAGGTCCCAATCAGCCCTTGTTCAGCATGTGCCATTTACGTCT 540
Qy 1060 GlnHisProHisValTyrSerProValIleGlnGlyAsnAlaAlaArgMetMetAlaProPro 1079
Db 541 CAGCATCCTCATGTATAGTCTGTATACAGGGTATGCTAGATGATGATGATGATGATGAT 600
Qy 1080 ThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGln 1099
Db 601 ACACAGCCCGAGCTGTTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAG 660
Qy 1100 ThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPhe 1119
Db 661 AGCATGCGCATGTAT----- 675
Qy 1120 TyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeu 1139
Db 676 -----GTTTCCAGGGCTCCCTTGTCTCAGCAGTATGCGCACCTTAACGCTACCTGT 726
Qy 1140 HisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGln 1159
Db 727 CACCCACATATCCACACCTCAGCCTTCCAGCTACCCCACTGGCAGCAGCAAGCCAA 786
Qy 1160 HisGlyGlySerHisPro-AlaProSer-ProValGlnHis-HisGlnHisGlnAla--- 1177
Db 787 AATGTGGGAAGTCACTCTGGCACCCAGTCCCTGGTCAGCACCCCATCAGCACCGGGCGG 846
Qy 1178 AlaGlnAlaLeuHisLeu 1183
Db 847 CCCAGGGTCTCCATCTG 864

RESULT 11
BQ716397
LOCUS
DEFINITION
AGENCY: 8477984 Lupski sympathetic_trunk Homo sapiens cDNA clone
IMAGE: 6196657 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 859)
AUTHORS
TITLE
NHL-MGC http://mgc.nci.nih.gov/
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
```

```
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13605 row: h column: 02
High quality sequence stop: 659.

FEATURES
Location/Qualifiers
1..859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6196657"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski sympathetic trunk"
/note="Vector: pCMV-Sport6 (Life Technologies); Site_1:
Noti; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCG-3' and
5'-GACTATCTCTAGATCGGCGGCGGCGCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 4,68e-66 Length: 859
Score: 1448.50 Matches: 277
Percent Similarity: 98.95% Conservative: 5
Best Local Similarity: 97.19% Mismatches: 2
Query Match: 20.81% Indels: 1
DB: 5 Gaps: 1

US-10-802-228-2 (1-1312) x BQ716397 (1-859)
Qy 555 CysProSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeuPro 574
Db 3 TGCCCATCTCTCTCTCTCGCCACCCTTCTCGTACCACTCAGGTCCCACTCTCTTCCA 62
Qy 575 ProArgAlaAlaThrProThrArgProProSerArgProProSerArgProSerArgPro 594
Db 63 CTTGGGCGAGCAGCCCTTACACGCGCGCCCTCCAGGCCCTCCGCGCCCATCCAGACCC 122
Qy 595 ProSerHisProSerAlaHisGlySerProAlaProValSerThrMetProLysArgMet 614
Db 123 CCGTCTCACCCCTCTGCTCATGGTTCTCCAGTCTCTGTCTACTATGCCCTAAACGCATG 182
Qy 615 SerSerGluGlyProProArgMetSerProLysAlaGlnArgHisProArgAsnHisArg 634
Db 183 TCTTCAGAGGCGCTCCCAAGGATGTCCCAAGGCCAGCCAGCATCTCTCGAAATCACAGA 242
Qy 635 ValSerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHisAsnProPro 654
Db 243 GTTTCTGTGGAGGGGTTCATATCCAGTGGCTAGAAATTTGTATCCCAACCCACCC 302
Qy 655 SerGluAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThrTrpSerSer 674
Db 303 AGTGAAGCAGCTACTCTCCAGTACAGGACCACTCCCTCGGGGGAACTGGTGTATCA 362
Qy 675 ValValSerGlyValProArgLeuSerProLysThrHisArgProArgSerProArgGln 694
Db 363 GTGGTCAGTGGGGTTCCAAGATTATCCCTAAAACCTATAGACCCAGGCTCTCCAGACAG 422
Qy 695 AsnSerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAlaGlyIleIle 714
```

Db 423 AACAGTATTGGAAATACCCAGTGGCCAGTCTTCTGCTTCTCCCAAGCTGTTATT 482

Qy 715 ProThrGluAlaValAlaMetProIleProAlaAlaSerProThrProAlaSerProAla 734

Db 483 CCAACTGAAGCTGTTGCCATGCTATTTCAGCTGCTATCTCTACGCTGCTAGTCTGCA 542

Qy 735 SerAsnArgAlaValThrProSerSerGluAlaLysAspSerArgGluGlnAspGlnArg 754

Db 543 TCGAACAGAGCTGTTACCCCTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATCAGAG 602

Qy 755 GlnAsnSerProAlaGlyAsnLysGluAsnLysGluAlaLysProAsnGluThrSerProSerPhe 774

Db 603 CAGAACTCTCTGCAGGGAATAAGAAATATTAACCCCAATGAACATCACTAGCTTC 662

Qy 775 SerLysAlaGluAsnLysGlyLysSerProValValSerGluHisArgLysGlnLysAsp 794

Db 663 TCAAAAGCTGAAACAAAGATATATCACCACTGTTGTTCTGAACATAGAAAACAGATTGAT 722

Qy 795 AspLeuLysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMet 814

Db 723 GATTTAAGAAATTTAAGATGATTATAGGTACAGCCCAAGTTCTACTTCTGAATCTATG 782

Qy 815 AspGlnLeuAsnLysAsnArgGluGlyGluLysSerArgAspLeu---lleLysAsp 833

Db 783 GATCACTACTAAACAAATAGAGGGGAGAGAAATCCAGAGATTGGATCAAGAC 842

Qy 834 LysIleGluProSer 838

Db 843 CAAATGAACCAAGT 857

RESULT 12

CA489164

LOCUS

DEFINITION AGENCOURT 10853363 MAPcL Homo sapiens cDNA clone IMAGE:6721472 5', mRNA linear EST 14-NOV-2002

ACCESSION CA489164

VERSION 1

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini, Hominidae; Homo.

1 (bases 1 to 889)

NTI-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Kristi A. Eglund, Ira Pastan

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution/LLNL at: found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

plate: MAM14282 row: k column: 08

High quality sequence stop: 683.

FEATURES

source

1. 889

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6721472"

/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCaP"

/lab_host="EMDHI08"

/clone_lib="NAPCL"

/note="Vector: pcMV-SPORT6; Site 1: EcoRV; Site 2: Not I; Subtracted with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dr. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg,

Bungkook Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

ORIGIN

Alignment Scores:

Pred. No.: 1.01e-65 Length: 889

Score: 1442.50 Matches: 284

Percent Similarity: 96.28% Conservative: 1

Best Local Similarity: 95.95% Mismatches: 8

Query Match: 20.72% Indels: 4

DB: 6 Gaps: 2

US-10-802-228-2 (1-1312) x CA489164 (1-889)

Qy 738 AlaValThrProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSer 757

Db 2 GCTGTTACCCCTTCTAGTAGGCTTAAAGATTCCAGGCTTCAAGATCAGAGGAGAACTCT 61

Qy 758 ProAlaGlyAsnLysGluAsnLysProAsnGluThrSerProSerPheSerLysAla 777

Db 62 CCTGCAGGGAATAAGAAATATTAACCCCAATGAACATCACTAGCTTCTCAAAAGCT 121

Qy 778 GlnAsnLysGlyLysSerProValValSerGluHisArgLysGlnLysAspLeuLys 797

Db 122 GAAACAAAGGTATATCACCACTGTTCTGAACATAGAAAACAGATTGATTTAAAG 181

Qy 798 LysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeu 817

Db 182 AAATTTAAGAATGATTTTAGGTTACAGCCCAAGTTCTACTTCTCAATCTATGATCAACTA 241

Qy 818 LeuAsnLysAsnArgGluGlyGluLysSerArgAspLeuLysAspLysLysGluPro 837

Db 242 CTAAACAAATAATAGAGGGGAGAGAAATCAAGAGATTGATCAAGACACAAAATTTGAACCA 301

Qy 838 SerAlaLysAspSerPheLysGluAsnSerSerSerAsnCysThrSerGlySerSerLys 857

Db 302 AGTGCTAAGGATCTTCTTCAATGAAATATAGCAGCAACTGTACCAAGTGGCAGCAGCAAG 361

Qy 858 ProAsnSerProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyPro 877

Db 362 CCGAATAGCCCCAGCAGCTTTCCCTTCAATCTAGTAACACGAGCAGCAAGAGGGGACCT 421

Qy 878 GluValThrSerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAsp 897

Db 422 GAGGTCACCTTCCCAAGGGGTTGAGACTTCCAGGCCAGCATGTAAACAAGAAAGACGAT 481

Qy 898 LysGluGluLysLysAspAlaGluGlnValArgLysSerThrLeuAsnProAsnAla 917

Db 482 AAGGAAGAGAAAGAAAGCAGCTGAGCAAGTTAGGAATTCACATTTGAATCCCAATGCA 541

Qy 918 LysGluPheAsnProArgSerPheSerGlnProLysProSerThrThrProThrSerPro 937

Db 542 AAGGAGTTCAACCCACGTTCTTCTCTAGCCAAAGCTTCTACTACCCCACTTCACT 601

Qy 938 ArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyr 957

Db 602 CGGCCTCAAGCACAACCTAGCCCATCTATGGTGGGTCTATCAACAGCCCACTTCCAGTTAT 661

Qy 958 ThrGlnProValCysPheAlaProAsnMetMetTyrProValProValSerProGlyVal 977

Db 662 ACTCAGCCCTGTTGTTTTCACCAATATATGATGATCCAGTCCAGTCCAGCCAGCGCTG 721

Qy 978 GlnProLeuTyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArg 997

Db 722 CAACCTTTATA-CCAATACCTATGACGCCCATGCCAGTGAATCAAGCCCAAGACATATAGA 780

Qy 998 Ala-----ValProAsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMet 1015

Db 781 GCAGGTAAGTACCAATATGCCCCACAGCGNCAAGCCAGCATCATCCAAAGGCCCATG 840

Qy 1016 MetHisProAlaSerAlaGlyPro---ProIleAlaAlaThrPro 1030


```

LOCUS      CD616048      830 bp      mRNA      linear      EST 12-JAN-2004
DEFINITION 56069370H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION  CD616048
VERSION    CD616048.1 GI:40264312
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE  1 (bases 1 to 830)
AUTHORS   Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE     Circular rapid amplification of cDNA ends for high-throughput
          extension cloning of partial genes
JOURNAL   Genomics 84 (1), 205-210 (2004)
PUBMED    15203218
COMMENT   Contact: Fu GK
          Incyte Genomics, Inc.
          3160 Porter Dr., Palo Alto, CA 94304, USA
          Tel: 6508454102
          Email: gfu@incyte.com.
FEATURES   Location/Qualifiers
            source
              1..830
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_lib="FLP"
              /note="Vector: pDrive Cloning Vector"
ORIGIN
Alignment Scores:
Pred. No.:      4,62e-64      Length:      830
Score:          1410.00      Matches:    264
Percent Similarity: 97.06%      Conservative: 0
Best Local Similarity: 97.06%      Mismatches: 7
Query Match:     20.26%      Indels:     2
DB:              6           Gaps:       0
US-10-802-228-2 (1-1312) x CD616048 (1-830)
QY      1040 TyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSer 1059
DB      829 TACAGTCCTCAGCAGGTGCCAAATACCCCTTTTCAGCATGTGCCACATATCATGCT 770
QY      1060 GlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProPro 1079
DB      769 CACATCCTCATGCTATAGTCTGTAAATACAGGGTAATGCTAGATGATGGCACCACCA 710
QY      1080 ThrHisAlaGln-ProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGI 1099
DB      709 ACACAGCCCGCCCTGGTTAGTATCTTCTCAGCAACTCAGTACGGGGCTCATGAGCA 650
QY      1099 nThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPh 1119
DB      649 GAGCGCATCGCATGTATGTCATGTGCCAA-TTACCATAACAAGAGAGACAAGCCCTCTTT 591
QY      1119 eTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLe 1139
DB      590 CTACTTTGCCATTTCACGGGCTCCCTTGCTCAGCAGTATGGCAGCCCTAACCGCTACCT 531
QY      1139 uHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGI 1159
DB      530 GCACCCACATACTCCACACCTTCAGCTTACCTTACCCCTGAGCAGCAGCAAGGCCA 471
QY      1159 nHisGlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGI 1179
DB      470 ACATGGTGAAGTCATCTGCACCCAGTCTCTGTTTCAGCACCGTCAGCACCAGCCGCCCA 411
QY      1179 nAlaLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaPr 1199
DB      410 GGCTCTCATCTGGCCAGTCCACAGCAGCAGTACGCCATTACCCCGGGGCTTGCC 351
QY      1199 oThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAl 1219

```

```

DB      350 AACTCCACCTCCATGACACCTGCTCCAAACGCGAGTCGCCACAGTAATAGTTTCCGAGC 291
QY      1219 aAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnPr 1239
DB      290 AGCACAAACAGAGACTGCTTTTACGATCCATCTCTCACGTTACGCGGGGTATACCAACCC 231
QY      1239 oProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisPr 1259
DB      230 ACCCCACATGGCCCCAGTACCTCAGGCTCATGTACAGTCAGGAATGGTTCCTTCATCC 171
QY      1259 oThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAl 1279
DB      170 AACTGCCCATGCGCCCAATGATGCTTAATGACGACAGCACCCCGGGTCCCGAGGCCG 111
QY      1279 aLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMe 1299
DB      110 CCTCGCTCAAAGTGCACTACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTATAT 51
QY      1299 tThrHisProSerValGlnAlaHisGlnGln 1310
DB      50 GACGCAACCTTCAGTACAAAGCCACCACCAACAG 17
RESULT 15
CX568338      865 bp      mRNA      linear      EST 12-JAN-2005
LOCUS      UI-M-180-cuv-e-17-0-UI.r1 NIH_BMAP_180 Mus musculus cDNA clone
DEFINITION IMAGE:30946480 5', mRNA sequence.
ACCESSION  CX568338
VERSION    CX568338.1 GI:57595367
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 865)
AUTHORS   NIH-MGC http://mgi.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James Lin University of Iowa
            cDNA library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
Seq primer: pyX-5.
            Location/Qualifiers
              1..865
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6"
              /db_xref="taxon:10090"
              /clone="IMAGE:30946480"
              /tissue_type="whole eye"
              /dev_stage="newborn (1,5,15 days)"
              /lab_host="DH10B (T1 phage resistant)"
              /clone_lib="NIH_BMAP_180"
              /notes="Organ: Eye; Vector: pyX-Asc; Site:1: EcoR I;
              Site 2: Not I; The library was constructed according
              Bonafido, Lennon and Soares, Genome Research, 6:791-806,
              1996. Denatured RNA was size fractionated on a 1% agarose
              gel. First strand cDNA synthesis was primed with oligo-dT
              size selected according to mRNA size fraction, ligated
              with EcoR I adaptor, digested with NotI and then cloned
              directionally into pyX-Asc vector. The library tag
              sequence located between the Not I site and the polyA tail
              is AATAATTACG. This library was created for the University

```


Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH).

ORIGIN

Alignment Scores:		
Pred. No.:	1.13e-63	Length:
Score:	1403.00	Matches:
Percent Similarity:	96.19%	Conservative:
Best Local Similarity:	92.04%	Mismatches:
Query Match:	20.16%	Indels:
DB:	8	Gaps:
		0
		2
		865
		266
		12
		10
		2

US-10-802-228-2 (1-1312) x CX568338 (1-865)

QY	440	AlaAsnGlnLeuAlaGluGluIleGlnSerSerAlaGlnTyrIysAlaArgValAla-Le	459
Db	2	GCAAACCACTTAGCAGAGAAATTTGA-TCCAGTCTCAGTACAAAGCTCGTGTGCGCCNT	60
QY	459	uGluAenAspArgSerGluGluGluLeuValThrAlaValGlnArgAsnSerSerG	479
Db	61	TCAGATATGATGCCGAGTGAGGAGAAAATAACAGCAGTCCAGAGAACTTCGACGTGA	120
QY	479	uArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAs	499
Db	121	CCGGAGGGGCATGCCCCAACACTAGGGACAATAATATATTCCTCTGGACAAAGAA	180
QY	499	nArgGluValIleSerTyrGlyArgGlnAsnSerProArgMetGlyGlnProG	519
Db	181	CAGAGAAGTCTTATCTCGGGAAAGTGGGAGACAGAGCTCACCGATGGCCGAGCTGG	240
QY	519	ySerGlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerG	539
Db	241	GCGAGCTCCATGCCGTCAAGAGCTGCTTCTCACACTTCAGATTCAACCCGAAACCTGG	300
QY	539	ySerAspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSe	559
Db	301	CTCAGACCAAAAGAGTAGTTAATGGAGGTGTTCCTCGCCATCGCTTGCCCATCTCTTC	360
QY	559	rSerArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaIaTh	579
Db	361	CTCTCGCCCACTTCTCGCTACCAAGTCAGGTCCCAACTCTTCTCACCTCGGCGACCA	420
QY	579	rProThrArgProProSerArgProProSerArgProSerArgProProSerHisProSe	599
Db	421	CCCTACAGGCCGCCCTCAGGCCCTCTGAGGCCCTCCAGACCCCGTCTCACCCCTTC	480
QY	599	rAlaHisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyPr	619
Db	481	TGCTCATGGTTCTCCAGCTCTCTGTCTCTACTATSCCTAAACGCATGTCTTCAGAGGAC	540
QY	619	oProArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyAr	639
Db	541	CCCAAGGATGTCCTCAAAAGGCACACGCCACCTTCGGAATCACAGAGTCTCTGCTGGAG	600
QY	639	gGlySerIleSerSerGlyLeuGluPheValSerHisAsnProProSerSerGluAlaIaTh	659
Db	601	AGGCTCCATGCTAGTGCCCTAGAAATTTGTATCCCAATCCCCCAAGTGAAGCAGCTGC	660
QY	659	rProProValAlaIaThThrSerProSerGlyGlyThrTyrProSerSerValValSerGlyVa	679
Db	661	TCCTCCAGTGGAAGACCACTCTCGAGGGGAAACGTGCTCTCAGTGGTCAGTGGGT	720
QY	679	lProArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAs	699
Db	721	TCCAAGGTATCTCCANAACTCAAGACCCAGGTCTCCAGGCGAGCAGCAGCATTTGAA	780
QY	699	nThrProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaVa	719
Db	781	CTCTCCAGCGGCCCTGTGCTGTCTTCTCCCAAGCTGGCATCATCTCTGCGAAGCCGT	840
QY	719	lAlaMetProfileProAlaIaSer	727

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 23, 2005, 06:06:48 ; Search time 513 Seconds

(without alignments)
4546.123 Million cell updates/sec

Title: US-10-802-228-2

Perfect score: 6961
Sequence: 1 MRSAAAPRSPAVATESRRF.....TAHFPYMTSPVQAHQOOL 1312

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xl
-Q/cgn2_1/USPTO_spool_p/US10802228/runat_21122005_110225_9729/app_query_faasta_1.1479
-DB=issued Patents NA -QFWT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10802228 @CGN 1.1 328 @runat_21122005_110225_9729 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6961	100.0	4481	3	US-09-041-886-18
2	6961	100.0	4481	3	US-09-648-281-1
3	6961	100.0	4481	3	US-09-707-919A-20
4	6961	100.0	4481	3	US-09-083-268-2
5	6961	100.0	4481	3	US-08-981-998A-2
6	6891.5	99.0	4484	3	US-09-949-016-4498
7	5606.5	80.5	3798	3	US-08-981-998A-4
8	1876	27.0	1257	3	US-09-648-281-11
9	1876	27.0	1257	3	US-09-083-268-4

10	1876	27.0	1257	3	US-08-981-998A-19	Sequence 19, Appl
11	1245.5	17.9	151088	3	US-09-949-016-16240	Sequence 16240, A
12	1012.5	14.5	623	3	US-09-043-303-5	Sequence 5, Appl
13	857	12.3	516	3	US-09-083-268-1	Sequence 1, Appl
14	857	12.3	516	3	US-08-981-998A-1	Sequence 1, Appl
15	787	11.3	459	3	US-09-707-919A-13	Sequence 13, Appl
16	643	9.2	355	3	US-09-043-303-1	Sequence 1, Appl
17	568	8.2	30690	3	US-09-914-286-1	Sequence 1, Appl
18	535.5	7.7	13987	2	US-08-804-227C-13	Sequence 13, Appl
19	535.5	7.7	44377	2	US-08-804-227C-7	Sequence 7, Appl
20	535.5	7.7	44377	2	US-08-804-198-1	Sequence 1, Appl
21	520.5	7.5	47677	3	US-09-949-002-668	Sequence 668, Appl
22	514	7.4	43280	2	US-08-804-227C-1	Sequence 1, Appl
23	501	7.2	34094	3	US-09-292-034-1	Sequence 1, Appl
24	498.5	7.2	4725	3	US-09-410-551B-24	Sequence 24, Appl
25	498.5	7.2	4725	3	US-09-940-316B-24	Sequence 24, Appl
26	497.5	7.1	33529	3	US-09-144-085-3	Sequence 3, Appl
27	497.5	7.1	50937	3	US-09-428-517-1	Sequence 1, Appl
28	494.5	7.1	8090	3	US-09-902-540-855	Sequence 855, Appl
29	493	7.1	31422	3	US-09-914-286-2	Sequence 2, Appl
30	490.5	7.0	80161	3	US-09-036-987A-1	Sequence 1, Appl
31	490.5	7.0	80161	3	US-09-370-700-1	Sequence 1, Appl
32	490.5	7.0	80161	3	US-09-603-207-1	Sequence 1, Appl
33	482	6.9	19598	3	US-09-902-540-1143	Sequence 1143, Appl
34	481.5	6.9	28958	2	US-08-258-261B-6	Sequence 6, Appl
35	481.5	6.9	28958	2	US-08-456-837-6	Sequence 6, Appl
36	481.5	6.9	28958	2	US-08-457-342-6	Sequence 6, Appl
37	481.5	6.9	28958	2	US-08-457-646A-6	Sequence 6, Appl
38	481.5	6.9	28958	2	US-08-458-076A-6	Sequence 6, Appl
39	481.5	6.9	28958	2	US-08-764-233A-4	Sequence 4, Appl
40	481.5	6.9	28958	2	US-08-457-335A-6	Sequence 6, Appl
41	481.5	6.9	28958	2	US-08-729-214-6	Sequence 6, Appl
42	481.5	6.9	28958	2	US-09-028-934-6	Sequence 6, Appl
43	481.5	6.9	47981	3	US-09-679-279-1	Sequence 1, Appl
44	481.5	6.9	49377	2	US-08-764-233A-1	Sequence 1, Appl
45	481	6.9	8438	2	US-07-945-283-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-041-886-18
; Sequence 18, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharoz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041,886
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LJ 2626
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949

! INFORMATION FOR SEQ ID NO: 18:
! SEQUENCE CHARACTERISTICS:
! LENGTH: 4481 base pairs
! TYPE: nucleic acid
! STRANDEDNESS: single
! TOPOLOGY: linear
! MOLECULE TYPE: DNA (genomic)
! FEATURE:
! NAME/KEY: CDS
! LOCATION: 163..4099
US-09-041-886-18

Alignment Scores:

Pred. No.: 1.33e-256 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-041-886-18 (1-4481)

Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 163 ATGGCGCTCAGCGCGCGCAGCTCTCGGAGTCCCGGTCGCGGTGGCCACCGAGTCTCGCCGCTTC 222
Qy 21 AlaAlaAlaArgTrrProGlyTrpArgSerLeuGlnArgProAlaArgArgSerGlyArg 40
Db 223 GCCGCAGCAGGTGGCCCGGGTGGCGTCCGCTCCAGCGCGCGCGCGGAGCGGGCGG 282
Qy 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProProGlyPro 60
Db 293 GCGCGCGGTGGCGCGCGCGCGCGAGCTATCCCTCCGCGCGCCCTCCCGCGCGCGCGCC 342
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaAlaSerAspCysPheGlySerAsn 80
Db 343 GGGCCCCCTCCCTCCGCGCAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAAC 402
Qy 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyGlyPro 100
Db 403 GGCAACGGCGCGCGCGGTTTCGGCGCGGCTCCCGCGGCTCTTGGTCTCGCGGGGCT 462
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaAlaProGlyAlaProProAlaAla 120
Db 463 CCGCGCGCTTCGTCGTCTCTCCCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 522
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 523 CCNACCGCGCGCTCCCGCTCGGCGCGCGCTCGCTCCCGCGCGCGCTTCGCGGCTCTCCTTG 582
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 583 GCGCGCGCGCTCCCGGCTGTCCCGCGCGGCTCGAGCGGCTATGGGCGCTCAC 642
Qy 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 643 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
Qy 181 GlnGlnGlnGlnGlnGlnProProProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 200
Db 703 CAG 762
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
Db 763 AGCGGCTTCTAGCGTCGCGCGCGCGCGCTTCGCGCTTCGCTCTCGGTCTCCTCG 822
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
Db 823 TCTCTCGGCGCAGCGCTCCCTCTCGGTGTCGCGCGAGCTCCGCGCGCGCGCGCGCGCG 882
Qy 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
Db 883 CTGGGCAGAGGTCGAAACAGTAACAGAGCTGCCTCAGTCTACGATTTCTTTTATGGA 942

Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
Db 943 ATCTATGCAAAATATGAGGATGGTTCATATACTTACATCAGTTGTTGGCTCCAAATGTCAA 1002
Qy 281 ValGlnValLysAsnGlyGlyIleTyrCluGlyValPheLysThrTyrSerProLysCys 300
Db 1003 GTACAAGTGAAAAATGGAGGTATATATGAAGGAGTTTTTAAACCTTACAGTCCGAAGTGT 1062
Qy 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
Db 1063 GATTGGTACTTGGTCCGCACATGAGAAAGATACAGATCCAGTTCGGGGCCGAAACGT 1122
Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValIleGlnPheLys 340
Db 1123 CAAGAAATAATGAGAGATATTTTGTCAAATGTTTTCAGACTTGTGTGGTACAGTTTAAA 1182
Qy 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db 1183 GATATGGACTCCAGTTATGCAAAAAGAGATCTTTTACTGACTCTGCTATCAGTGCCTAAA 1242
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla 380
Db 1243 GTGAATGGCGAACAACAAGAGAGGACCTGGAGGCCCTGGGATGAGGTGAACACTCACGCC 1302
Qy 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMet 400
Db 1303 AATCAGGAACTTGAGGCTTTTGGAAAAATGACGTATCTAATGATGGGATCCCAATGATATG 1362
Qy 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
Db 1363 TTTTCGATATATGAAGAAAAATTTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db 1423 TATACAGTCCCTTAGAAGAGATACTCAGAGAAGATTTTTTAAACCGGAAGCAAGGCA 1482
Qy 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db 1483 AACCACTTAGCAGAAGAAATTTAGTCAAGTCCCAAGTCAAAAGCTCAGTGGCCCTGAA 1542
Qy 461 AsnAspArgArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db 1543 AATCATATAGGAGTGAGGAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGAGTGAACGT 1602
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
Db 1603 GAGGGGCACAGCATAAACACTTAGGGAATAATAATATATCTCTCTGGACAAAGAAATAGA 1662
Qy 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1663 GAAGTCATATCTCGGGAAGTGGGAGACAGAAATTCACCGGTATGGGCCAGCCTGGATCG 1722
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTTCAGATTTCAACCCCGAAATCTCGTTCA 1782
Qy 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCCCTTCCCATCTCTCTCTCTCT 1842
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1843 CGCCCACTTCTCGCTACCACTCAGGTCCCAACTCTCTTCCACTCTCGGCGCAGCCACCCCT 1902
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db 1903 ACACGGCGCGCTCCAGCGCCCTCCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1963 CATGGTCTCCAGCTCTCTGCTCTACTATGCCTAAACGATGTCTTCAGAAAGGCGCTCCA 2022

Qy		621	ArgMetSerProIysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly	640
Db		2023	AGGATGTCTCCCAAGGCCACAGCACACTCTTCGAAATCAGAGATTTCTGTGGAGGGGT	2082
Qy		641	SerIleSerSerGlyLeuGluPheValSerHisAsnProSerProSerGluAlaAlaThrPro	660
Db		2083	TCCATATTCCAGTGGCCTAGAATTGTGATCCACAACCACCCAGTGAAGCAGCTACTCT	2142
Qy		661	ProValAlaArgThrSerProSerSerGlyGlyThrTipSerSerValValSerGlyValPro	680
Db		2143	CCAGTAGCAGGACCAGTCCCTCGSGGGGAACGTGGTGTCATCAGTGGTCAGTGGGGTTCCA	2202
Qy		681	ArgLeuSerProIysThrHisArgProProArgSerProArgGlnAsnSerIleGlyAsnThr	700
Db		2203	AGATTATCCCCATAAACCTCATAGACCAGGTCTCCCAGACAGAAACAGTATTGGAAATACC	2262
Qy		701	ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla	720
Db		2263	CCCAGTGGGCAGTTCTTGCTCTCCCCAAGCTGGTATTATTCCAAGTGAAGCTGTGCC	2322
Qy		721	MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr	740
Db		2323	ATGCCTATTCCAGCTGCATCTCTACGCCTGCTAGTCTGCTCATCGAACAGAGCTGTACC	2382
Qy		741	ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly	760
Db		2383	CCTTCTAGTCAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCAGAATCTCTCTGCAGGG	2442
Qy		761	AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys	780
Db		2443	AATAAGAAATAATTAAACCAATGAACATCATCCTAGCTTCTCAAAGCTGAAAACAAA	2502
Qy		781	GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys	800
Db		2503	GGTATATCACCACTGGTTCTGAACATAGAAAAACAGATTGATGATTAAAGAAATTAAG	2562
Qy		801	AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys	820
Db		2563	AATGATTTTAGGTTACAGCCAAAGTTCTVACTTCTGAATCTATGGATCAACTACTATAACAAA	2622
Qy		821	AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys	840
Db		2623	AATAGAGAGGGAGAAAATCAAGAGATTTGATCAAGACAAAAATTGAACCAAGTGCTAAG	2682
Qy		841	AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer	860
Db		2683	GATTCCTTTCATTGMAAATAGCAGCGACACTGTACCAAGTGGCAGCAAGCCGATAGC	2742
Qy		861	ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr	880
Db		2743	CCCAGCATTTCCCTTCAATACTTTAGTAACACGGAGCACAAAGGGGACCTGAGGTCAC	2802
Qy		881	SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu	900
Db		2803	TCCCAAGGGGTTCAGACTTCCAGCCCAGCATGTAAACAGAGAAAGACGATAAGGAAGAG	2862
Qy		901	LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe	920
Db		2863	AAGAAAGACGACGTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGGAGTTTC	2922
Qy		921	AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln	940
Db		2923	AAACCAAGTTCCTTCTCTCAGCCAAAGCCTTCTACTACCCCACTTCACTCGGCCCTCAA	2982
Qy		941	AlaGlnProSerProSerMetMetValGlyHisGlnGlnProThrProValTyrThrGlnPro	960
Db		2983	GCACAACCTAGCCCATCTATGGTGGGTTCATCAACAGGCAACTCCAGCTTTTATCTACGSC	3042
Qy		961	ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu	980
Db		3043	GTITGTTTTGCACCAATAATGATGTATCCAGTCCAGTGAAGCCAGCGGTGCACACCTTTA	3102
Qy		981	TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro	1000

Db	3103	TACCCAAATACCTATGACGCCCATGCCAGTGAATCAAGCCAGACATATAGACAGTTACCA	3162
Qy	1001	AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMethHisProAlaSer	1020
Db	3163	AATATGCCCAACAGCGGCAAGACACAGCATCATCATGAGTGCCATGATGCACCCAGCGCTCA	3222
Qy	1021	AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr	1040
Db	3223	GCACGGGCCCCACCGAATGCGAGCCACCCACCAAGCTTACTCTCACGCAATATGTTGGCTAC	3282
Qy	1041	SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln	1060
Db	3283	AGTCCTCAGCAGTTCCCAATCAGCCCTTGTTTCAGCATGTGCCACATTATCAGTCTCAG	3342
Qy	1061	HisProHisValTyrSerProValIleGlnGlnYAsnAlaArgMetMetAlaProProThr	1080
Db	3343	CATCCTCATGCTATATAGTCTCTGTAATACAGGGTAATGCTAGAAATGATGCGCACCAACCA	3402
Qy	1081	HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr	1100
Db	3403	CAGCCCGAGCCTGGTTTAGTATCTTCTTCAGCAACTCAGTACGGGCTCATGACGACAGC	3462
Qy	1101	HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr	1120
Db	3463	CATGCGATGTATGATGTCTCCAAATTACCATACAACAGGAGACAAGCCCTCTTTCTCTAC	3522
Qy	1121	PheAlaIleSerThrGlySerIleAlaGlnTyrAlaHisProAsnAlaThrLeuHis	1140
Db	3523	TTTGCCCAATTTCCACGGGCTCCCTTGCTCAGCAGTATGCGCACCTAACCGTCACTTCAC	3582
Qy	1141	ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis	1160
Db	3583	CCACATACTCCACACCCCTCAGCCTTCAGCTACCCCACTGGACAGCAGCAAGGCCAACAT	3642
Qy	1161	GlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAla	1180
Db	3643	GGTGGAAAGTCATCCTGCACCCAGTCTCTGTTCACGACCATCAGCACCGCGCCCAAGCT	3702
Qy	1181	LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr	1200
Db	3703	CTCCATCTGGCCAGTCCACAGCAGCAGTACGCCATTTACACCGGGGGCTTGGCCCACT	3762
Qy	1201	ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla	1220
Db	3763	CCACCTCCCATGACACCTGCCTCCAAACAGCAGTCGCGCACAGAAATAGTTTCCCGACAGCA	3822
Qy	1221	GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro	1240
Db	3823	CAACAGACTGTCTTTACGATCCATCCTTCCTCAGCTTCAGCGGGCGGTATACCAACCCACC	3882
Qy	1241	HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr	1260
Db	3883	CACATGCCCCACGTACCTCAGGCTCATGTACAGTCAGGAATGGTTCTCTCTCATCCAAT	3942
Qy	1261	AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu	1280
Db	3943	GCCCATCGGCAATATGATGCTAATAGACACACAGCCACCCGGCGGTATCCCAAGCGCCCTC	4002
Qy	1281	AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr	1300
Db	4003	GCTCAAAGTGCACTACAGCCCATTCAGTCTCGACAACACAGGCATTTCCCTCATATGACG	4062
Qy	1301	HisProSerValGlnAlaHisHisGlnGlnGlnLeu	1312
Db	4063	CACCCTTTCAAGTACAGGCCACCCACCAACAGCAGTTG	4098

RESULT 2
US-09-648-281-1
; Sequence 1, Application US/09648281
; Patent No. 6515197
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M.

; TITLE OF INVENTION: Transgenic Animal Model of
; TITLE OF INVENTION: Neurodegenerative Disease and Methods of Use
; FILE REFERENCE: P-CE 4336
; CURRENT APPLICATION NUMBER: US/09/648,281
; CURRENT FILING DATE: 2000-08-24

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 4481

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (163)...(4099)

US-09-648-281-1

Alignment Scores:

Pred. No.: 1,33e-256 Length: 4481

Score: 6961.00 Matches: 1312

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-648-281-1 (1-4481)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
DB 163 ATGCGCTCAGCGCGCGAGCTCTCGAGTCCCGGGTGGCCACCGAGTCTCGCGCTTC 222
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
DB 223 GCCGAGCAGGTGGCGCGGGTCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGG 282
QY 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
DB 283 GCGCGCGTGGCG 342
QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
DB 343 GSCCGCCCTCCCTCCGCGCAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAAC 402
QY 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyGlyPro 100
DB 403 GCGAAACGGCGCGCGCGCGTTCGCGCGCGCTCCGCGCGCTCTCTGGTCTCGCGCGCGCT 462
QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
DB 463 CCGCGCGCGCTTCGTGTGTCTTCTCCCGCTCGCGCGCGCGCGCGCGCGCGCGCG 522
QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB 523 CCAACCGCGCGCTCCCGCTCGCGCGCGCGTGGTCCCGCGCGCTTCGCGGTCTCTCTTG 582
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB 583 GCGCGCGCGCTTCGCGGTCTCCCGCGCGCGCGTTCGCGCGCGCTGTATGGCGCGCTCACC 642
QY 161 MetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 643 ATGTCTGTGAAGCCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762
QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerSerSerSer 220
DB 763 AGCGCGCTTTCAGCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 822
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
DB 823 TCCTCGCGCACCGCTCCCTCTCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 882

QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
DB 883 CTGGCGCAGAGGTGCAAAACAGTAACAAAGGAGCTGCCTCAGTCTACGATTTCTTTTGGATGGA 942
QY 261 IleTyrAlaAsnMetArgMetValHileLeuThrSerValValGlySerLysCysGlu 280
DB 943 ATCTATGCAAAATATGAGGATGGTTTATATCTTACATCAGTTGTGGCTCCAAATGTGAA 1002
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB 1003 GTACAAGTGAATAATGAGGATATATATGAAGAGTGTTTTAAACCTTACAGTCCCAAGTGT 1062
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
DB 1063 GATTTGGTACTTGTATGCGGCACATGAGAAAGTACAGAAATCCAGTTCGGGGCCGAAACGT 1122
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
DB 1123 GAAGAAATAATGAGAGAGTATTTTGTTCAAATGTTTTCAGACTTTGTGTGGTACAGTTTAAA 1182
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB 1183 GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTCTAAA 1242
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla 380
DB 1243 GTGAATGCGCAACAAAGAGAGGACCTGGAGCCCTGGGATGCGAGTGAATCAGACCC 1302
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMet 400
DB 1303 AATGAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGATG 1362
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB 1363 TTTTCGATATAATGAAGAAAAATATGTTGTGTAGTGTCTAGTATGATGACAGTTTATCTTCG 1422
QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB 1423 TATACAGTGCCCTTAGAAAGAGATAACTCAGAAGAAATTTTAAACGGGAAGCAAGGCA 1482
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1483 AACCAAGTTAGCAGAAAGAAATTTAGTCAAGTCCAGTACAAAGCTCAGTGGCCCTGGA 1542
QY 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB 1543 AATGATGATAGAGTGAAGAAAGAAATAACACAGAGTTTACAGAGAAATTCACAGTGAACGT 1602
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500
DB 1603 GAGGGCAGACAGATAAAACACTAGGAAAAATATAATATATCTCTCTGGACAAAGAAATAGA 1662
QY 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB 1663 GAAGTCAATATCTCGGGAAGTGGAGAGACAGAAATTTACCGCGTATGCGGCGACCGCTGGA 1722
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTTCAGATTTCAACCCCGAATCTCTGGTTCA 1782
QY 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
DB 1783 GACCAAGAGTAGTTTAAATGGAGGTGTTCCCTGGCCATCGCCTTGCCTTCCCTCTCTCTCT 1842
QY 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
DB 1843 CGCCCAACCTTCTCGCTACCACTAGTCCCAACTCTCTTCCACCTCGGCGACCAACCCCT 1902
QY 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
DB 1903 ACACGCGCGCCCTCAGAGCCCGCCCTCGGCGCATTCAGACCCCGCTCTCACCCCTCTGCT 1962


```
; Sequence 20, Application US/09707919A
; Patent No. 6623927
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: Method for detection of human spinocerebellar ataxia 2
; TITLE OF INVENTION: gene variants
; FILE REFERENCE: US 443
; CURRENT APPLICATION NUMBER: US/09/707,919A
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-707-919A-20

Alignment Scores:
Pred. No.: 1,33e-256 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-707-919A-20 (1-4481)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaAlaThrGluSerArgArgPhe 20
DB 163 ATGGCTCAGCGCGCGAGCTCTCGGAGTCCGCGGTGGCCAGCTCTCGCGCTTC 222
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
DB 223 GCCGAGCAGCGTGGCGCGGTGGCGTCCGTCAGCGCGCGCGCGCGAGCGCGG 282
QY 41 GlyGlyGlyValAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
DB 283 GCGCGCGTGGCGCGCGCGCGGACGATATCCCTCCGCGCGCGCGCGCGCGCG 342
QY 61 GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
DB 343 GSCCCCTCCCTCCGCGCAGAGCTCGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 402
QY 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyGlyPro 100
DB 403 GGCACAGCGCGCGCGCGTTCGCGCGCGCTCCGCGCGCTCTGGTCTCGCGCGGCT 462
QY 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
DB 463 CCCCCCTTCGTCGTCTCTCTCCCTCCGCGCAGCGCGCGCGCGCGCGCGCG 522
QY 121 ProThrArgAlaSerProLeuGlyAlaAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB 523 CCAACCGCGCTCCCGCTCGCGCGCGTGGGTCCCGCGCGGTTCGCGGTCTCTT 582
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB 583 GCGCGCGCGGTCCCGGTTCGCGCGCGCGTCCGCGCGTCCGAGCGGTATGGGCCCTCACC 642
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 643 ATGTCGTGAAGCCCAACGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762
QY 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220
DB 763 AGCGGCTTCTAGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 822
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaAlaThrSerGlyGlyArgProGly 240
DB 1962
```

```
823 TCCTCGGCCACGCGCTCCCTCCTCGTGTGTCGCGCGACCTCCGCGCGGAGGCCCGC 882
241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
893 CTGGCAGAGGTGCAACAGACTAAACAAGGACTGCCTCAGTCTACGATTTCTTTGATGGA 942
261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
943 ATCTATGCAAAATATGAGATGCTTATATGAGGAGTTCATATGAGGAGTTCATGAGGAGT 1002
281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
1003 GTACAAGTGAAGGAGGAGTATATGAGGAGTTCATGAGGAGTTCATGAGGAGTTCATGAGGAGT 1062
301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
1063 GATTGGTACTTGGTGGCGCACATGAGAAAGATGAGAAATCCAGTTCGGGCGGCAACGT 1122
321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
1123 GAAGAAATAATGAGAGATATTTTGTCAAAATGTTTCAAGACTTGTGTGATACAGTTTAAA 1182
341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA 1242
361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380
1243 GTGAATGGCGAACACAAAGAGAGAGACTGGAGCGCTGGGATGAGGTAACTACAGCC 1302
381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
1303 AATGAGAACTTGAGGCTTTGGAATAATGACGTATCTAATGATGGGATCCCAATGATG 1362
401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
1363 TTTCTGATATATGAGAAATAATGAGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
1423 TATACAGTGCCTTAGAAAGAGATTAATCAGAAAGATTTTAAACGGGAAGCAAGGCA 1482
441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
1483 AACCAAGTATGAGAAAGAAATGAGTCAAGTCCAGTCAAAAGCTCGAGTGGCGCTGGA 1542
461 AsnAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
1543 AATGATATAGGAGTGAAGAAAGAAATACACAGCAGTTTCCAGAGAAATTCAGTGAACGT 1602
481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500
1603 GAGGGGCACAGCATTAACACTAGGGGAAATAAATAATATATCTCTCTGGAACAAGAAATAGA 1662
501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
1663 GAAGTCATATCTGGGAAAGTGGGAGACAGAAATTCACCGCTATGGGCGACCGCTGGATCG 1722
521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTTCAGATTTTCAACCCGAAATTCCTGGTTCA 1782
541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560
1783 GACCAAGAGATGATTAATGAGGAGTTCCTCTGGCCATCGCCTTCCCATCTCTCTCTCTCT 1842
561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
1843 CGCCCACTTCTCGCTACCACTCAGTCCAGGTCCTTCTTCCACTCTCTCCAGCTCGGGCAGCCACCT 1902
581 ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla 600
1903 ACAGCGCGCGCGCTCCAGCGCGCGCGCTCGCGCGCATCCAGACCGCGCTCTACCGCTCTGCT 1962
```

QY	601	HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro	620
DB	1963	CATGGTCTCCAGCTCCTGCTCTACTATGCTTAAACGATGCTTTCAGAGGGCTCCA	2022
QY	621	ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly	640
DB	2023	AGGATGTCCTCCAAAGGCCAGCGACATCTCTGAAATACAGAGATTCTCTGGAGGGGT	2082
QY	641	SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerSerGluAlaAlaThrPro	660
DB	2083	TCCATATCCAGTGGCCCTAGAAATTTGATCCCAACCCAGTCGAGCAGCTACTCT	2142
QY	661	ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro	680
DB	2143	CCAGTAGCAAGGACCAAGTCCCTCGGGGGAAACGTGTCATCAGTGTGCTGAGTGGGTCCA	2202
QY	681	ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr	700
DB	2203	AGATTATCCCTTAAAACTCATAGCCAGGCTCTCCAGACAGAACAGTATTGGAATACC	2262
QY	701	ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla	720
DB	2263	CCAGTGGGCCAGTCTTCTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGC	2322
QY	721	MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr	740
DB	2323	ATGCCTATTCCAGCTGCATCTCTAGCGCTGCTAGTCTCTGCATCGAACAGAGCTGTATCC	2382
QY	741	ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly	760
DB	2383	CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGAGCAACTCTCTCGAGGG	2442
QY	761	AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys	780
DB	2443	AATPAAGAAATATTAAACCATGAAACATCACTAGCTTCTCAAAAGCTGAAACAAA	2502
QY	781	GlyIleSerProValValSerGluHisArgLysGlnIleAspLeuLysLysPheLys	800
DB	2503	GGTATATCACCACTGTTCTGAACATAGAAAACAGATTGATGATTTAAAGAAATTTAAG	2562
QY	801	AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys	820
DB	2563	AATGATTTTAGTTACAGCCAAGTCTTACTTCTGAATCTATGGATCAACTACTATAACAA	2622
QY	821	AsnArgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys	840
DB	2623	AATAGAGAGGGAGAAATCAAGAGATTGATCAAGACAAATTTGAACCAAGTGTCTAAG	2682
QY	841	AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer	860
DB	2683	GATTTCTTCAATGAAATAGCAGCAGCAACTGTACCACTGGCAGCAGCAAGCCGAATAGC	2742
QY	861	ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr	880
DB	2743	CCCAGCATTTTCCCTTCAATACTTAGTAACACCGAGCACAAGAGGGGACCTGAGGTCACT	2802
QY	881	SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu	900
DB	2803	TCCCAAGGGGTTTCACTTCCAGCCCGCATGTAAACAGAGAAAGCATAGAGAGAG	2862
QY	901	LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe	920
DB	2863	AAGAAAGACGACGCTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGAGATTC	2922
QY	921	AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln	940
DB	2923	AACCCAGCTTCTCTCAGCCAAAGCTTCTACTACCCCAACTTCACTCGCGCTCAA	2982
QY	941	AlaGlnProSerProSerMetValGlyHisGlnProThrProValThrThrGlnPro	960
DB	2983	GCACAACTAGCCCATCTATGGTGGGTCAACAGCCAACTCCAGTTTATCTACGCTCT	3042

RESULT 4

QY	961	ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu	980
DB	3043	GTTTGTTTGGCACCATAATATGATGATCCAGTCCAGGTAGCCAGCGGTGCAACCTTTTA	3102
QY	981	TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro	1000
DB	3103	TACCCAAATACCTATCAGCCCATGTCAGTGAATCAAGCCACACATATAGAGCAGTACCA	3162
QY	1001	AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer	1020
DB	3163	AATATGCCCCAACAGCGGCAAGACCATCATCAGAGTGCCATGATGACCCAGCGTCA	3222
QY	1021	AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr	1040
DB	3223	GCAGCGGGCCCCACCATTTGCAGCCACCCACAGCTTACTCCAGCAATATGTTGGCTTAC	3282
QY	1041	SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln	1060
DB	3283	AGTCTCTAGCAGCTTCCCAAAATCAGCCCTTGTTCAGCATGTCCACATTTATCAGTCTCAG	3342
QY	1061	HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr	1080
DB	3343	CATCTCTATGTCATAGTCTCTGTAATACAGGGTAATGCTAGAAATGATGCGCACCAACA	3402
QY	1081	HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr	1100
DB	3403	CACGCCACCCCTGGTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGAGC	3462
QY	1101	HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr	1120
DB	3463	CATGCGATGATGTCATGTCCTCCAAATTTACCATACACAGAGACAGCCCTTCTTTCTAC	3522
QY	1121	PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis	1140
DB	3523	TTTGGCATTTCCAGCGGCTCCCTTGTCTCAGCAGTATGCGCACCTAACGCTACCTCGCAC	3582
QY	1141	ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis	1160
DB	3583	CCACATACTCCACACCTCAGCTTCCAGTACCCCACTGGAGCAGCAGCAAGCCCAACAT	3642
QY	1161	GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla	1180
DB	3643	GGTGGAGTCATCTGTCACCCAGTCTCTGTTACAGCACCATCAGCCAGCGCCCGCAGGCT	3702
QY	1181	LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr	1200
DB	3703	CTCCATCTGGCCAGTCCACAGCAGCAGTCAGGCCATTTACACCGGGGCTTGGCCCAACT	3762
QY	1201	ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla	1220
DB	3763	CCACCTCTCATGACACTGCTCTCCAAACAGCGAGTCGCCACAGAAATAGTTTCCAGCAGCA	3822
QY	1221	GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro	1240
DB	3823	CAACAGACTGCTTTACAGATCCATCTTCTCAGCTTACGCGCGGTATACCAACCCACCC	3882
QY	1241	HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr	1260
DB	3883	CACATGGCCCCACGTACTCTCAGGCTCATGTACAGTCAGGAATGGTTCTCTCTCATCCAAT	3942
QY	1261	AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu	1280
DB	3943	GCCCATGCGCCAAATGATGCTAATGACACACAGCCACCCCGCGGTCCCCAGCGCGCCCTC	4002
QY	1281	AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr	1300
DB	4003	GCTCAAGTGCACTACAGCCCATTCAGTCTCGAACAAACAGCGCATTTCCCTATATGACG	4062
QY	1301	HisProSerValGlnAlaHisHisGlnGlnLeu	1312
DB	4063	CACCTTTCAGTACAAAGCCCCACCAACCAACAGCAGCTTG	4098

US-09-083-268-2
 ; Sequence 2, Application US/09083268
 ; Patent No. 6673535
 ; GENERAL INFORMATION:
 ; APPLICANT: Pulst, Stefan M
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
 ; TITLE OF INVENTION: ATAXIA-2 AND PRODUCTS RELATED THERETO
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Muefing, Raasch & Gebhardt, P.A.
 ; STREET: 119 No. 6673535th Fourth Street
 ; CITY: Minneapolis
 ; STATE: Minnesota
 ; COUNTRY: USA
 ; ZIP: 55401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/083,268
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/727,084
 ; FILING DATE: 08-OCT-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McCormack, Myra H
 ; REGISTRATION NUMBER: 36,602
 ; REFERENCE/DOCKET NUMBER: 232.00010101
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 612/305-1220
 ; TELEFAX: 612/305-1228
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4481 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: both
 ; TOPOLOGY: both
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 163..4101
 ; US-09-083-268-2

Alignment Scores:
 Pred. No.: 1,33e-256 Length: 4481
 Score: 6961.00 Matches: 1312
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-083-268-2 (1-4481)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
 DB 163 ATGCGCTCAGCGCGCGAGCTCCTCGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTC 222
 QY 21 AlaAlaAlaArgTTPProGlyTTPArgSerLeuGlnArgProAlaArgSerGlyArg 40
 DB 223 GCCGAGCAGGTGGCCCGCGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCGCGG 282
 QY 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
 DB 283 GCGCGGGTGGCGCGCGCGCGCGGATCCCTCGCGCGCGCGCGCGCGCGCGCGCG 342
 QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
 DB 343 GCGCGCGCGCTCCCGCGCGAGAGCTCGCGCTCCCTCCGCTCAGACTGTTTTGGTAGCAAC 402
 QY 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyGlyPro 100

DB 403 GGCAACGGCGCGCGCGCTTTCCGCCCGCGCTTCCGCCCGCGCTTCTGGTCTCGCGCGGCGCT 462
 QY 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
 DB 463 CCCCGCGCGCTTTCGTCGTCTCTTCTCCCTTCGCCAGCCCGCGCGCGCTTCCGCCCGCG 522
 QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProAlaArgSerGlyValSerLeu 140
 DB 523 CCAACCGCGCGCTCCCGCTCGCGCGCGCTCGCTCCCGCGCGCTTCCGCCCGCTTCTCTTG 582
 QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
 DB 583 GCGCGCGCGCTTCCCGCTTCCCGCGCGCGCTTCCCGCGCGCTTCCCGCGCGCTTCCAC 642
 QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
 DB 643 ATGTCGCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
 QY 181 Gln 200
 DB 703 CAG 762
 QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerProSerSerSerSerSerSer 220
 DB 763 AGCGCGCTTCTAGCGTCCCGCGCGCGCGCTTCCGCCCGCTTCCGCCCGCTTCTCTCG 822
 QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGlyArgProGly 240
 DB 823 TCCTCGCGCACGGCTCCCTCTCGTGGTCCGGCGAGCTCCGGCGCGCGCGCGCGCGCG 882
 QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
 DB 883 CTGGCGCAGAGGTGCAAAACAGTAACAAAGAGTGGCTCAGTCTAGTCTCTTTGATGGA 942
 QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
 DB 943 ATCTATGCAAAATATGAGGATGGTTCATATCTTACATCAGTTGTGGCTTCCAAATGTCAA 1002
 QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
 DB 1003 GTACAAGTGAAATGGAGGTATATATGAAGAGATTTTAAACCTTACAGTCCGAAGTGT 1062
 QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerSerGlyProLysArg 320
 DB 1063 GATTTGGTACTTGTATGCGCACATGAGAAAGTACAGAAATCCAGTTCGGGGCCGAAACGT 1122
 QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
 DB 1123 GAAGAAATAATGGAGAGTATTTTGTCAAATGTTTCAAGTCTTGTGTGTACAGTTTAA 1182
 QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
 DB 1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA 1242
 QY 361 ValAsnGlyGluHisLysGlyLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380
 DB 1243 GTGAATGGCGAACAACAAGAGAGAGCCTGGAGCGCTGGAGTGAAGTCAACAGCC 1302
 QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
 DB 1303 AATCAGGAACTTTGAGGCTTTGCAAAATGACGTATCTAATGATGGATGCCCAATGATATG 1362
 QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
 DB 1363 TTTTCGATATATGAAGAAATATATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
 QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
 DB 1423 TATCAGTGCCTTTAGAAAGAGATAACTCAGAAGAAATTTTAAACCGGGAAGCAAGGCA 1482
 QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460

1483 AACCAAGTTAGCAGAGAAATTGAGTCAAGTCCAGTACAAAGCTCGAGTGCCTGGNA 1542
461 AsnAspAspArgSerGluGluGlySerThrAlaValGlnArgAsnSerSerGluArg 480
1543 AATGATGATAGGAGTGAGGAGAGAAATAACACAGCAGATTCCAGAGAAATCCAGTGAACGT 1602
481 GluGlyHisSerIleAsnThrArgGluAsnLysThrIleProProGlyGlnArgAsnArg 500
1603 GAGGGGCACAGCATAAACACTAGGGAATAATAATATATTCCTCGTGAGCAAGAAATAGA 1662
501 GluValIleSerThrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
1663 GAAGTCATATCTCGGGAAGTGGAGACAGAAATTCACCGCGTATGGCCGAGCTGGATCG 1722
521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTCAACCCGGAATTCCTGGTTCA 1782
541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
1783 GACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCCTTGGCCATCTCTCTCTCTCT 1842
561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
1843 CGCCCACTTCTCGCTACCACTGAGTCCCACTCTCTTCCACTCGGCGAGCCACCCCT 1902
581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
1903 ACAGGCGCCCTCCAGGCCCCCTCGGGCCATCCAGACCCCGCTCACCCCTCTGTCT 1962
601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
1963 CATGGTTCTCCAGCTCTGCTCTACTATGCTTAAACGCGATGCTTTTCAGAGGGGCTCCA 2022
621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
2023 AGGATGTCCCAAGGCCCGACACATCTCTGAAATCAACAGATTTCTCTCTGGAGGGGT 2082
641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
2083 TCCATATCCAGTGCCTAGAAATTTGATCCCAACCCAGTGAAGCAGCTACTCTCT 2142
661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680
2143 CCAGTAGCAAGGACCAAGTCTCGGGGGAACGTGTGTCATCAGTGTGCTGAGTGGGTCCA 2202
681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
2203 AGATTATCCCTAAAACCTCATAGCCAGGCTCTCCAGACAGAACAGTATTTGAAATACC 2262
701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
2263 CCAGTGGGCGCAGTTCTTGCTTCTCCCAAGCTGTTATTTCCAACTGAGCTGTGCT 2322
721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
2323 ATGCTTATTCAGCTGCATCTCTACGCTGTCTGCTGATCGATCGAACAGAGCTGTATCC 2382
741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
2383 CCTTCTAGTAGGCTAAAGATTCAGGCTTCAAGATCAGAGGCGAGAACTCTCTCTGAGGG 2442
761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
2443 AATAAGAAATATTAACCCATGAACATCACCTAGCTTCTCNAAGCTGAACACAA 2502
781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
2503 GGTATATCACAGTTGTTCTGAACATAGAAAAACAGATTGATGATTTAAAGAAATTAAG 2562
801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
2563 AATGATTTTAGTTACAGCCCAAGTTCTTCTGTAATCTATGATCAACTACTATAACAA 2622

821 AsnArgGluGlyGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
2623 AATAGAGAGGAGAGAAATCAAGAGATTGTCAAAAGACAAATTTGAACCAAGTCTTAAG 2682
841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
2683 GATTTCTTTCATTTGAATAATAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAATAGC 2742
861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
2743 CCCAGCATTTCCCTTCAATACTTAGTAACAGGAGCACAAAGAGGGACCTCAGGTCACT 2802
881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGlyLysAspAspLysGluGlu 900
2803 TCCCAAGGGGTTTCAGACTTCCAGCCCGCATGTAAACAGAGAAAGACATTAAGGAAGAG 2862
901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
2863 AAGAAAGACGCACTGAGCAAGTTAGGAATCAACATTGAATCCCAATGCAAGGAGTTTC 2922
921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
2923 AACCCAGCTGCTCTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACCTCGGCTCAA 2982
941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
2983 GCACAACCTAGCCCATCTATGTGGGTCTATCAACAGCCCAACTCCAGTTTATACTCAGCCT 3042
961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
3043 GTTTGTTTGGACCAAAATATGATGATCCAGTCCCAAGTCCAGCCGCGGTGCAACCTTTA 3102
981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
3103 TACCCAATACCTATGAGGCCCATGTCAGTGAATCAAGCCCAAGACATATAGAGCAGTACA 3162
1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
3163 AATATGCCCAACAGCGGCAAGCAGCATCATCAGAGTGCCATGATGACCCAGCGTCA 3222
1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
3223 GAGCGGGGCCCAACCATTCAGCCACCCACAGCTTACTCCAGCAATATGTTGCTTAC 3282
1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
3283 AGTCTCAGCAGTTCCTCAATCAGCCCTTGTTCAGCATGTGCCATATTCAGTCTCAG 3342
1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
3343 CATCTCATGTCTATAGTCTCTGTAATACAGGGTAAATGTAGATGATGACCAACCA 3402
1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
3403 CACGCCACCCCTGGTTTGTATCTCTTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462
1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
3463 CATGCGATGTATGTCATGTCCCAATTTACCATACAAAGAGACAAAGCCCTTCTTCTAC 3522
1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
3523 TTTGCCATTTCCAGGGCTCCCTTGTCTCAGCAGTATGCGCACCTAACGCTACCTCGCAC 3582
1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
3583 CCACATATCCACACCTCAGCCTTCCAGTACCCCACTGGACAGCAGCAAGCCCAACAT 3642
1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
3643 GGTGGAAGTCTCTCTGCACTCCAGTCTGTGTTTCAGCACCATCAGCACCGCCCGCCAGCT 3702

Db 943 ATCTATGCAAAATATGAGGATGGTTCATATACATCACTGTTGGCTCCAAATGTGAA 1002
Qy 281 ValGlnValIleAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db 1003 GTACAAAGTGAATAATGAGGTATATATGAAGGAGTGTATTAACCTACAGTCCGAGTGT 1062
Qy 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
Db 1063 GATTTGGTACTTGAATCCGACATGAGAAAGTACAGAAATCCAGTTCGGGGCGGAAACGT 1122
Qy 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
Db 1123 GAAGAAATATGAGAGATATTTTGTCAATGTTACAGCTTTGTGTTGTGACAGTTTAA 1182
Qy 341 AspMetAspSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db 1183 GATATGGACTCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTCTATCAGTGTCTAAA 1242
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380
Db 1243 GTCAATGGCGAAACACAAAGAGAGAGACCTGGAGCCCTGGGATGAGGTGAACCTCACAGCC 1302
Qy 381 AsnGluGluLeuAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
Db 1303 AATGAGGAACCTTGAGGCTTTGGAAATGACGTATCTAATGGATGGGATCCCAATGATATG 1362
Qy 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
Db 1363 TTTGATATATGAGAAATATATGTTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db 1423 TATACAGTGCCTTAGAAAGAGATAACTCAGAGAATTTTTTAAACCGGAAGCAAGGGCA 1482
Qy 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db 1483 AACCAAGTTACGAGAAATATGATCAAGTCCAGGTACAAAGCTCGAGTGGCCCTGGAA 1542
Qy 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db 1543 AATGATGATAGGATGAGGAGAAATAATACACAGCAGTTTCAGAGAAATTCAGTGAACGT 1602
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500
Db 1603 GAGGGGCACAGCATAACACTAGGAAATAAATATATCTCTCTGGACAAAGAAATAGA 1662
Qy 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1663 GAAGTCTATATCCTGGGGAATGGGAGACAGAAATTCACCGGTATGGGCCAGCCTGGATCG 1722
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1723 GGCTCCATGTCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATCTGGTTCA 1782
Qy 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560
Db 1783 GACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCTTGGCCCATCTCTCTCTCT 1842
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1843 CGCCCACTTCTCGCTACCAAGTCAAGTCCCAACTCTCTCCACCTCGGGCAGCCACCCCT 1902
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db 1903 ACAGGCCGCCCTCCAGGCCCTCCGGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1963 CATGGTTCTCAGCTCCTGCTCTACTATGCTTAAACGCATGCTTTCAGAGGGGCTCCA 2022
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Db 2023 AGGATGTCCCAAGGGCCAGCGACATCTCGAAATCACAGAGTTTCTGCTGGGAGGGGT 2082

Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
Db 2083 TCCATATTCAGTGGCTAGAAATTTGTATCCCAACACCCAGTGAAGCAGCTACTCTCT 2142
Qy 661 ProValAlaAlaThrSerProSerGlyGlyThrTyrTrpSerSerValValSerGlyValPro 680
Db 2143 CCAGTAGCAAGAGCCAGTCCCTCGGGGGAACGTGGTCATCAGTGGTCACTGGGGTTC 2202
Qy 681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db 2203 AGATTATCCCTTAAACCTCATAGACCCAGGTCTCCAGACAGACAGTATTCGAAATACC 2262
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
Db 2263 CCCAGTGGGCCAGTTCTTCTCTCCCAAGCTGGTATTTATTTCCAACTCAAGCTGTTC 2322
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
Db 2323 ATGCTTATTTCCAGCTGCATCTCTACGCTGTAGTCTCTGCATCGAACAGAGCTGTACC 2382
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnArgGlnAsnSerProAlaGly 760
Db 2383 CCTTCTAGTAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCGCAGGG 2442
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
Db 2443 AATAAAGAAAAATATTAACCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAAAA 2502
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLys 800
Db 2503 GGTATATCACCAAGTTGTTCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2562
Qy 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db 2563 ATGATTTTAGGTTACAGCCAGTTCTTACTTCTGAATCTATGATCAACTACTATAACAAA 2622
Qy 821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
Db 2623 AATAGAGAGGGAGAAAAAATCAAGAGATTTCATCAAAAGACAAAAATGAAACCAAGTCTAAG 2682
Qy 841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Db 2683 GATTTCTTTCATTTGAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAAATAGC 2742
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db 2743 CCAGCATTTCCCTTCAATACTTATTAGACCGAGCACAAGAGGGGACCTGAGGTCACT 2802
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
Db 2803 TCCCAAGGGGTTTCAGACTTCCAGCCCGCATGTAAAAACAAGAGAAAGACGATAAGAAAGAG 2862
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 2863 AAGAAAGACGACGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAGGAGTTC 2922
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2923 AACCCAGTTCTTCTCTCAGCCAAAGCTTCTTACTACCCCAACTTCCACCTCGGCCTCAA 2982
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db 2983 GCACAACTTAGCCCATCTATGTTGGGTTCATCAACAGCCAACTCCAGTTTATCTACGCT 3042
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db 3043 GTTTGTTTTCACCAAAATATGATGTATCCAGTCCCGAGTGGCCGAGCGCTGCAACCTTTA 3102
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 3103 TACCAATATCTATGACGCCCATGCCCAGTGAATCAAGCCAAAGACATATAGAGCAGTACCA 3162


```
QY 1001 AenMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
DB 3163 AATATGCCCCAACAGCGCAAGCAGCATCATCAGAGTGCCATGATCACCACCGCTCA 3222
QY 1021 AlaaAlaGlyProProIleAlaaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
DB 3223 GCAGCGGGCCCAACCATGTCAGCCACCCACACAGCTTACTCCACCAATATGTTGCCATC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
DB 3283 AGTCCTCAGCAGTTCCTCCAAATCAGCCCTTGTTTCAGCATGTGCCACATTCAGTCTCAG 3342
QY 1061 HisProHisValTyrSerProValIleGlnGlnGlyAsnAlaArgMetMetAlaProProThr 1080
DB 3343 CATCTCTCATGTCTATAGTCCTGTAATACAGGGTAATGCTAGAAATGATGGCACCAACA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyValHisGlnGlnThr 1100
DB 3403 CACGCCACCGCTGGTTAGTATCTTCTTCAGCACTCAGTACGGGGCTCATGAGCAGACG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
DB 3463 CATCGATGATGATGATGCTCCCAATACCATACAAAGGAGACAAAGCCCTCTTTCTAC 3522
QY 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
DB 3523 TTTGTCATTTTCACCGGGCTCCCTTCTCAGCAGTATGGCACCTTAAGCTACCTCGCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
DB 3583 CCACATATCTCCACACCTCAGCTCCTCAGCTACCCCTGAGACAGACAGAAAGCAACAT 3642
QY 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
DB 3643 GGTGGAAGTCATCTCTGACCCAGTCTCTGTTACGACCATCAGCACAGCGCCCGCCAGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
DB 3703 CTCCATCTGGCCAGTCCACAGCAGCAGTACAGCCATTTACACAGCGGGGCTTGCGCAACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
DB 3763 CCACCTCTCATGACACCTGCTCCAAACGCGAGTCCGACAGAAATGTTTCCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
DB 3823 CAACGAGCTGTCTTTACGATCCATCTCTCAGCTTACGGCGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3883 CACATGGGCCCACTACCTCAGCTCATGTACAGTCAGGAATGGTTCTCTCTCATCCAAC 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
DB 3943 GCCCATGGCCCAATGATGCTAATGACGACACAGCCACCCGCGGGTCCCCAGCGCCCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrAlaHisPheProTyrMetThr 1300
DB 4003 GCTCAAGTGCACTACGCCCATTCAGTCTCGACACAGCGCATTTCCCTATATGAGC 4062
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
DB 4063 CACCCCTTCAGTACAAAGCCCAACCAACACAGCAGTTG 4098
```

RESULT 6

```
US-09-949-016-4498
; Sequence 4498, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
```

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4498
; LENGTH: 4484
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4498

Alignment Scores:
Pred. No.: 5,87e-254 Length: 4484
Score: 6891.50 Matches: 1309
Percent Similarity: 99.70% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 2
Query Match: 99.00% Indels: 3
DB: 3 Gaps: 1
```

US-10-802-228-2 (1-1312) x US-09-949-016-4498 (1-4484)

```
QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
DB 163 ATGGCTCAGCGCGCGAGCTCTCCGAGTCCCGGGTGGCCAGGCTCTCGCG- TTC 221
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
DB 222 GCCCAGCAGCTGGCGGGTGGCTCGCTCCAGCGCGCGCGCGCGGAGCGGCGG 281
QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
DB 282 GCGCGCGTGGCGGGCGCGGACCGTATCCCTCCGCGCGCTTCTCTCCCGCGGCC 341
QY 61 GlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
DB 342 GCCCGCTCTCTCCGCGCAGCTCCCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 401
QY 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100
DB 402 GGCACCGCGCGCGGCTTTCGCGCGGCTCCCGCGGCTCTTGGTCTCGCGGGCT 461
QY 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
DB 462 CCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 521
QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB 522 CCAACCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 581
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB 582 GCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 641
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 642 ATGTCGCTGAAGCCCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 701
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
DB 702 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 761
QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerValSerSer 220
DB 762 AGCGGCTTCTAGCGTCCCGCGCGCGCGCTTCCGCGCTCTCTCTCTCTCTCTCT 821
QY 221 SerSerAlaThrAlaProSerSerValAlaAlaThrSerGlyGlyGlyArgProGly 240
DB 822 TCCTCGGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 881
```

QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
DB |||||
DB 882 CTGGGCAGAGGTCGAACACAGTAAACAGACTGCCTCAGTCTACGATTTCTTTTATGGA 941
QY |||||
DB 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
DB |||||
DB 942 ATCTATGCAAAATATGAGGATGTTTCAATATCTATCATCTGTTGGTCTCCAAATGTGAA 1001
QY |||||
DB 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB |||||
DB 1002 GTACAAGTGAAATATGAGGATATATGAGGATTTTAAACCTTACAGTCCGAAGTGT 1061
QY |||||
DB 301 AspLeuValLeuAspAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
DB |||||
DB 1062 GATTTGGTACTTGTATGCCCATGAGAAAGTACAGAAATCCAGTTCGGGGCGGAAACGT 1121
QY |||||
DB 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
DB |||||
DB 1122 GAAGAAATATGAGAGTATTTTGTTCAAATGTTTCAAGCTTTGTGTGATCAGTTTAA 1181
QY |||||
DB 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB |||||
DB 1182 GATATGGACTCCAGTATGCANAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAA 1241
QY |||||
DB 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrPheAspAlaGlyGluLeuThrAla 380
DB |||||
DB 1242 GTCAATGGGGCAACACAAAGAGAAGGACCTGGAGCCCTGGGATCGAGGTGAATCTCACAGCC 1301
QY |||||
DB 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
DB |||||
DB 1302 AATGAGAACTTGAGGCTTTGGGAAATGACGTATCTAATGAGTGGATCCCAATGATATG 1361
QY |||||
DB 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB |||||
DB 1362 TTTTCGATATATGAAGAAATATGTTGTGTAGTCTACGTATGATAGCAGTTTACTTCG 1421
QY |||||
DB 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB |||||
DB 1422 TATACAGTGCCTTAGAAAGAGATAACTCAGAAGAAATTTTAAACGGGAAGCAAGGGCA 1481
QY |||||
DB 441 AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrIleAlaArgValAlaLeuGlu 460
DB |||||
DB 1482 AACCACTTAGCAAGAAATGAGTCAAGTGCCCGAGTACAAAGCTCGAGTGCCCTGGAA 1541
QY |||||
DB 461 AsnAspAspArgSerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB |||||
DB 1542 AATGATATAGGAGTGAAGAGAAATAATACACAGCAGTTTCAGAGAAATTCAGTGNACGT 1601
QY |||||
DB 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500
DB |||||
DB 1602 GAGGGCCACAGCATAAACACTAGGAAATAAATAATATATCTCTCTGGACAAAGAAATAGA 1661
QY |||||
DB 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB |||||
DB 1662 GAAGTCATATCTCTGGGAGTGGGAGACAGAAATTCACCGGTATGGCCGAGCCCTGGATCG 1721
QY |||||
DB 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB |||||
DB 1722 GGTCTCATGCCATCAAGATCCATCTTCACACTTCAGATTTCAACCCGGAATTCGTGTCA 1781
QY |||||
DB 541 AspGlnArgValValAsnGlyGly----ValProTyrProSerProCysProSerProSe 559
DB |||||
DB 1782 GACCAAGAGTAGTTAATGAGGCAAGTGTTCCTGGCCATCGCTTGGCCATCTCCTTC 1841
QY |||||
DB 559 rSerArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaIle 579
DB |||||
DB 1842 CTCTCGCCCACTTCTCGTACAGTCCAGTCCCAACTCTCTTCCACCTCGGGCAGCCAC 1901
QY |||||
DB 579 rProThrArgProProSerArgProProSerArgProSerArgProProSerHisProSe 599
DB |||||
DB 1902 CCCTACAGGCGCCCTTCAGGCCCCCTCGGGCCATCCAGACCCCCCGTCTCACCCCTC 1961

QY 599 rAlaHisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyPr 619
DB |||||
DB 1962 TGCTCATGGTCTCAGCTCCTGTCTTACTATGCTTAACGCATGTCTTTCAGAGGGCC 2021
QY |||||
DB 619 oProArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyAr 639
DB |||||
DB 2022 TCCAAAGATGTCCCAAGGCCAGCGACATCTCGAAATCACAGAGTTCTGTCTGGCAG 2081
QY |||||
DB 639 gGlySerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaTh 659
DB |||||
DB 2082 GGGTTCCATATCCAGTGGCCTAGAAATTTGTATCCCAACACCACCCAGTGAAGCAGCTAC 2141
QY |||||
DB 659 rProProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyVa 679
DB |||||
DB 2142 TCCTCCAGTAGCAAGGACAGTCCCTCGGGGGAAAGCTGGTTCATCAGTGGTGGGT 2201
QY |||||
DB 679 lProArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAs 699
DB |||||
DB 2202 TCCAAAGATTTATCCCTTAAACTCATAGACCAGGTCTCCACAGACAGACAGTATTGGAAA 2261
QY |||||
DB 699 nThrProSerGlyProValLeuAlaSerProGlnAlaGlyIlelleProThrGluAlaVa 719
DB |||||
DB 2262 TACCCCGAGTGGCCAGTTCTGTCTCTCCCAAGCTGGTATTATTCCAACCTGAAGCTGT 2321
QY |||||
DB 719 lAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnAtGAlaVa 739
DB |||||
DB 2322 TGCCATGCTTATTTCCAGCTGCATCTCTACGCTCTAGTCTCTGCATCGAACAGAGCTGT 2381
QY |||||
DB 739 lThrProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAl 759
DB |||||
DB 2382 TACCCCTTCTAGTGAGGCTTAAAGATTCCAGGCTTCAAGATCAGAGCGCAACTCTCTGC 2441
QY |||||
DB 759 aGlyAsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAs 779
DB |||||
DB 2442 AGCGAATAAGAAAAATATTAAACCCCAATGAACATCACTAGCTTCTCAAAAGCTGAAAA 2501
QY |||||
DB 779 nLysGlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPh 799
DB |||||
DB 2502 CAAAGGTATATCACCAGTTGTTTCTGAACATAGAAAAACAGATTGATGATTTTAAAGAAAT 2561
QY |||||
DB 799 eLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAs 819
DB |||||
DB 2562 TAAAGATGATTTAGTTAGTTACAGCCAGTTCTACTTCTGAATCTATGGATCAACTATAA 2621
QY |||||
DB 819 nLysAsnArgGluGlyGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAl 839
DB |||||
DB 2622 CAAAAATAGAGAGGAGAAAAATCAAGAGATTGATCAAAAGACAAAAATGAAACCAAGTGC 2681
QY |||||
DB 839 aLysAspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAs 859
DB |||||
DB 2682 TAAAGATTTCTTTTCAATTGAAATAGCAGCAGCAACTGTACAGTGGCAGCAGCAAGCCGAA 2741
QY |||||
DB 859 nSerProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluVa 879
DB |||||
DB 2742 TACCCCGCAGATTTTCCCTTCAATCTAGTAAACAGCGAGCACAAGAGGGGACCTGAGGT 2801
QY |||||
DB 879 lThrSerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGly 899
DB |||||
DB 2802 CACTTCCCAAGGGTTTCAGACTTCCAGCTCCAGCCAGCATGTAAACAAAGAAAAAGACGATAAGGA 2861
QY |||||
DB 899 uGluLysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGly 919
DB |||||
DB 2862 AGAGAAAGAAAGACGAGCTGAGCAAGTTAGGAAATCAACATTTGAAATCCCAATGCAAGGA 2921
QY |||||
DB 919 uPheAsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgPr 939
DB |||||
DB 2922 GTTCAACCCAGCTTCTCTCAGCCAAAGGCTTCTACTACCCCAACTTCCACTCGGCC 2981
QY |||||
DB 939 oGlnAlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGly 959
DB |||||
DB 2982 TCAAGCACAACTAGCCCATCTATGTGGGTCTATCAACAGCCAACTCCAGTTTATACTCA 3041
QY |||||
DB 959 nProValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnPr 979

```
Db 3042 GCCTGTTGTTGGACCAATATGATGATCCAGTCAGCCAGCCAGGCGTCAACC 3101
QY 979 oLeuTyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaVa 999
Db 3102 TTATATACCAATACCTATGACGCCCATGCGAGTGAATCAAGCAAGACATATAGAGCAGT 3161
QY 999 lProAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAl 1019
Db 3162 ACCAATATGCCCAACACGCGCAAGACAGATCATCAGAGTGGCATGATGCACCCAGC 3221
QY 1019 aSerAlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAl 1039
Db 3222 GTCAGAGCGGGCCCAACCAATGTCAGCCACCCACAGCTTACTCCAGCAATATGTTGC 3281
QY 1039 aTyrSerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSe 1059
Db 3282 CTACAGTCCTCAGCAGTTCCCAAAATCAGCCCTTGTTCAGCATGTGCCACATATATCAGTC 3341
QY 1059 rGlnHisProHisValTyrSerProValIleGlnGlnGlnAlaArgMetMetAlaProPr 1079
Db 3342 TCAGATCCTCATGCTATAGTCCTGTAATACAGGGTAATGCTAGAAATGATGGCACCCACC 3401
QY 1079 oThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGl 1099
Db 3402 AACACAGCGCCAGCCTGGTTAGTATCTCTTCAGCAACTCAGTACGGGCTCATGAGCA 3461
QY 1099 nThrHisAlaMetTyrAlaCysProGlyLeuProTyrAsnLysGlnThrSerProSerPh 1119
Db 3462 GACGATCGCATGTATGATGTCATGTCCTCCAAATTAACATACAAAGGAGACAGCCCTTCCT 3521
QY 1119 eTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrIle 1139
Db 3522 CTACTTTGCCATTCTCCAGCGGCTCCTTGTCTCAGCAGTATGGCACCCCTTAACGCTACGCT 3581
QY 1139 uHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGl 1159
Db 3582 GCACCCACATATCTCACACCTTCAGCTTACGCTACCCCTCATGACAGCAGCAAGGCA 3641
QY 1159 nHisGlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaAlaGl 1179
Db 3642 ACATGGTGAAGTCATCCTGCACCCAGTCTCTTTCAGCACCATCAGCACCGCCGCCCA 3701
QY 1179 nAlaLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaPr 1199
Db 3702 GGCTCTCCATCTGGCCAGTCCACAGCAGCAGTCAGCCATTTACACCGCGGGCTTGGGCC 3761
QY 1199 oThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAl 1219
Db 3762 AACTCCACCTCCATGACACCTGCTCCAAACGCGAGTCGCCACAGATAGTTCCTCCAGC 3821
QY 1219 aAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnPr 1239
Db 3822 AGCACAAACAGACTGCTTTACGATCCATCTTCTCAGCTTACGCGGGGTATACCAACC 3881
QY 1239 oProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisPr 1259
Db 3882 ACCCCACATGGGCCACGTACTCCTCAGGCTCATGTACAGTCAGGAATGGTTCTCTCATCC 3941
QY 1259 oThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAl 1279
Db 3942 AACTGCCATGGCCAAATGATGCTAATGACGACACAGCCACCGCGGTGCCAGGCCGC 4001
QY 1279 aLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMe 1299
Db 4002 CCTGCTCAAAAGTGACATACAGCCCATTCAGTCTCGACACAGCGCATTTCCCTATAT 4061
QY 1299 tThrHisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db 4062 GACGACACCTTCAGTACAGGCCACCCACCAACAGCAGATTG 4101
```

RESULT 7

US-08-981-998A-4

```
; Sequence 4, Application US/08981998A
; Patent No. 6844431
; GENERAL INFORMATION:
; APPLICANT: PULST, STEFAN M.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCERECELLAR
; ATAXIA-2 AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MUEITING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: 55401
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,998A
; FILING DATE: 11-May-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 97/42314
; FILING DATE: 08-MAY-1997
; APPLICATION NUMBER: US 08/727,084
; FILING DATE: 08-OCT-1996
; APPLICATION NUMBER: US 60/022,207
; FILING DATE: 19-JUL-1996
; APPLICATION NUMBER: US 60/017,388
; FILING DATE: 08-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MUEITING, ANN M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 232.00010120
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-305-1217
; TELEFAX: 612-305-1228
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3798 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 50..3457
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-981-998A-4
Alignment Scores:
Pred. No.: 4,02e-205 Length: 3798
Score: 5606.50 Matches: 1063
Percent Similarity: 94.20% Conservative: 41
Best Local Similarity: 90.70% Mismatches: 37
Query Match: 80.54% Indels: 31
DB: 3 Gaps: 8
US-10-802-228-2 (1-1312) x US-08-981-998A-4 (1-3798)
QY 148 ProArgProAlaCysGluProValTyrGlyProLeuThrMetSerLeuLysProGlnGln 167
Db 11 CCCCAGCGGGGTGGAGCGCGGTGTATGGCGGCTCACCATGTCTGTAAGCCGAGCGG 70
QY 168 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 187
Db 71 CAG----- 73
QY 188 ProProProAlaAlaAlaAsnValArgLysProGlyGlySerGlyLeuLeuAlaSerPro 207
Db 74 ---CCGCCCGCGCCCGCCACTGGCGCGCAAGCCCGCGGC---GGCCTGCTCTCTGCGGCC 127
```


Db 2282 CGTTCTTCTCAGCCAAAGCCTTCTACTACCCCAAGCTCACCTCGCGCTCAAGACAA 2341
Qy 943 ProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnProValCys 962
Db 2342 CCCAGCCCATCTATGTTGGTTCATCAGCAGCAGCTCCAGTGTATACCTCAGCCTGTGTC 2401
Qy 963 PheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeuTyrPro 982
Db 2402 TTGCGACCCAAATATGATGTATCCGCTCCAGTGAGCCCGGCGGTACAACCTTTATACCA 2461
Qy 983 IleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAla-----ValPro 1000
Db 2462 ATACCTATGAGCCCATGCTGTGAACCAAGCAAGACATATAGAGAGTAAAGTACCA 2521
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
Db 2522 AATATGCCCCAACAGCAGACCAACATCATCAAGCACCATGATGACCCAGCCTCC 2581
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 2582 GCGGAGGCGCCACCATCGTAGCCACCCCGCGCTTACTCCACTCAGTAGCTTGCCTAC 2641
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 2642 AGCCCTCAGCATTTCCCAATCAGCCTTTGGTCCAGCATGTGCCGCATTATCAGTCTCAG 2701
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 2702 CATCTCATGTGTACAGCTCTGTCTACAGTAATGCCAGGATATGGCCACCACCA 2761
Qy 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyValHisGluGlnThr 1100
Db 2762 CATGCTCAGCCTGGTTTAGTGTCTTCTCAGCTGCTCAGTTCGGGGCTCAGCAGCAGC 2821
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 2822 CACGCCATGTATGATGTATGCCAAATACCATACAACAAGGAGACAAAGCCCTTCTTCTAC 2881
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 2882 TTTGGCATTTCCACCGGCTCCCTCGCTCAGCAGTATGACATCTCTAATGCGCCCTGCT 2941
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Db 2942 CCACATATCCCATCTCAGCCTTCGGCCACTCCACCGGACAGCAGCAAAAGCCAGCAT 3001
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAla 1180
Db 3002 GGTGGAAGTCAACCTGCAACCCAGTCTGTTCAGCACCATCAGCACAGGCTGCCAGGCT 3061
Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3062 CTTTCAITGGCCAGTCCACAGCAGCTGGGCCATTTATCATGCGGGGCTGGCACCAAC 3121
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3122 CCACCTTCCATGACACCTGCCTCTAATACACAGTCTCCACAGAGCAGTTCCTCCAGCAG 3181
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3182 CAACAGACAGTCTTCCACCATCCACCTCTCATGTTCAGCGCGCATACACACCCACCC 3241
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3242 CACATGGCCCAAGTACCTCAGGCTCATGTACAGTCAGGAATGTTCTTCATCCAAC 3301
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3302 GCCCATGGCCCAATGATGTAATGACGACACAGCCACCC---GGTCCCAAGCCGCCCTC 3358
Qy 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrAlaHisPheProTyrMetThr 1300

Db 3359 GCTCAAAGTGCACTACAGCCCATTTCCAGTTTCGACAAACAGCGCATTTCCCTTATATGACG 3418
Qy 1301 HisProSerValGlnAlaHisHisGlnGlnGlnLeu 1312
Db 3419 CACCTTCAGTACAAGCCCAACACACACAGCAGTTG 3454
RESULT 8
US-09-648-281-11
; Sequence 11, Application US/09648281
; Patent No. 6515197
; GENERAL INFORMATION:
; APPLICANT: Pulist, Stefan M.
; TITLE OF INVENTION: Transgenic Animal Model of
; FILE OF INVENTION: Neurodegenerative Disease and Methods of Use
; FILE REFERENCE: P-CE 4336
; CURRENT APPLICATION NUMBER: US/09/648,281
; CURRENT FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1257)
US-09-648-281-11
Alignment Scores:
Pred. No.: 1,3e-63 Length: 1257
Score: 1876.00 Matches: 365
Percent Similarity: 89.77% Conservative: 21
Best Local Similarity: 84.88% Mismatches: 18
Query Match: 26.95% Indels: 26
Gaps: 5
US-10-802-228-2 (1-1312) x US-09-648-281-11 (1-1257)
Qy 157 GlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGln 176
Db 8 GGGCGCTCACATGTCGTGAAGCCGACGCGGAG----- 43
Qy 177 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
Db 44 -----CCGCCCGCGCGCCGCGCGC 67
Qy 197 LysProGlyGlySerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSer 216
Db 68 AAGCCCGCGCGC---GGCTGCTCTCGTCCGCGCGCGCGCGCGC---TCGCGCGCG 121
Qy 217 SerValSerSerSerAlaThrAlaProSerValValAlaAlaThrSer----- 234
Db 122 GTGACCTCGCTTCGTGTGTCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 181
Qy 235 ---GlyGlyGlyArgProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGln 253
Db 182 GCGGGCGCGCGCGCTCCCGGCTCGGCGAGAGTTCGGAACAGTAGCAAAAGGACTGCTCAG 241
Qy 254 SerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSer 273
Db 242 CCTACGATTTCTTTTGTGGAATCTATGCAAAACGTGAGGATGGTTTCATATCTTACGTC 301
Qy 274 ValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPhe 293
Db 302 GTTGTGGATCAAAATGTGAAGTACAGTGAAACCGAGGACATATATCAAGGAGTTTTT 361
Qy 294 LysThrTyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGlu 313
Db 362 AAAACATACAGTCTCCTAAGTGTGACTTGGTACTTGTGCTGCACATGAGAAAAGTACAGAA 421
Qy 314 SerSerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp 333
Db 422 TCCAGTTCCGGGCCCAACCGTGAAGAAATAATGGAGAGTGTGTTTGTTCATAATGCTCAGAC 481

```
334 PheValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThr 353
482 TTCGTTGTGGTACAGTTTAAAGATACAGACTCCAGTTATGACCGGAGAGATGCTTTTACT 541
354 AspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTyr 373
542 GACTCTGCTCTCAGCGCAAGAGTGAATGTGTAGCACAAGGAGAGACCTGGAGCCCTGG 601
374 AspAlaGlyGluLeuThrAlaAsnGluLeuGluAlaLeuGluAsnAspValSerAsn 393
602 GATGACGAGGAGCTCAGCGCAGGAGGAGCTGGAG---CTGGAGNATGATGTCTAAT 558
394 GlyTyrAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThr 413
659 CGATGGGACCCCAATGACATGTTTCATATATATGATGAGAGAAATATGCTGTGTGCCACA 718
414 TyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPhe 433
719 TATGATAGCAGTTTATCTTCATATACGGTTCTTTAGAAAGGAGCAACTCAGAAGAAATT 778
434 LeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyr 453
779 CTTAAACGGGAGCAGGGGCAACCACTGTTAGCAGAAGAAATTGAATCCAGTCTCAGTAC 838
454 LysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluGluLysTyrThrAlaVal 473
839 AAAGCTCGTGTGCGCTTCAGAAATGATGACCGGAGTGAGAGAGAAATAACACAGCAGTC 898
474 GlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIle 493
899 CAGAGAAACTGCAGTACCGGGAGGGGCGATGGCCCAACACTAGGAGACAATAATATATT 958
494 ProProGlyGlnArgAsnArgGluValIleSerTyrGlySerGlyArgGlnAsnSerPro 513
959 CTTCTGGACAAAGAAACAGAGAAAGTCTTATCTCTGGGAAAGTGGGAGACAGACTCACCA 1018
514 ArgMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAsp 533
1019 CGATGGGCGAGCTGGGCGAGGCTCCATGCGCTCAAGAGCTGCTTCTCACACTTCAGAT 1078
534 PheAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTyrProSer 553
1079 TTCAACCGCAACGCTGGCTCAGACCAAGAGAGTAGTTAATGGAGGTGTTCCTGGCCATCG 1138
554 ProCysProSerProSerArgProProSerArgTyrGlnSerGlyProAsnSerLeu 573
1139 CCTTGCCCATCTCTCTCTCCGCCACCTTCTCGCTACCAAGTCAAGTCCCAACTCTCTT 1198
574 ProProArgAlaAlaThrProThrArgPro 583
1199 CCACCTCGGGCAGCCACCCCTACACGGCT 1228
```

RESULT 9

```
US-09-083-268-4
; Sequence 4, Application US/09083268
; Patent No. 6673535
; GENERAL INFORMATION:
; APPLICANT: Pulist, Stefan M
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
; TITLE OF INVENTION: ATAXIA-2 AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Muetting, Raasch & Gebhardt, P.A.
; STREET: 119 No. 6673535th Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/083.268

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/727,084

FILING DATE: 08-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: McCormack, Myra H

REGISTRATION NUMBER: 36,602

REFERENCE/DOCKET NUMBER: 232.00010101

TELECOMMUNICATION INFORMATION:

TELEPHONE: 612/305-1220

TELEFAX: 612/305-1228

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1257 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 2..1255

US-09-083-268-4

Alignment Scores:

Pred. No.:	1,3e-63	Length:	1257
Score:	1876.00	Matches:	365
Percent Similarity:	89.77%	Conservative:	21
Best Local Similarity:	84.88%	Mismatches:	18
Query Match:	26.95%	Indels:	26
DB:	3	Gaps:	5

US-10-802-228-2 (1-1312) x US-09-083-268-4 (1-1257)

```
QY 157 GlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGln 176
DB 8 GGGCGCGCTCACCATGTGCTGAAGCGCGCAGCGCAG----- 43
QY 177 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
DB 44 -----CCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 67
QY 197 LysProGlyGlySerGlyLeuLeuAlaSerProAlaAlaProSerProSerSer 216
DB 68 AAGCCCGGCGGC---GGCCTGCTCTCGTCCCGCGCGCGCGCGCGCGCGCGCG 121
QY 217 SerValSerSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSer 234
DB 122 GTGACCTCGGCTTCCGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 181
QY 235 ---GlyGlyGlyArgProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGln 253
DB 182 GCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 241
QY 254 SerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSer 273
DB 242 CCTACGATTTCTTTGATGGAATCTATGCAACAGCTGAGGATGTTTCATATACGTCA 301
QY 274 ValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPhe 293
DB 302 GTTGTGGATCGAAATGTGAAGTACAAAGTGAAGAAACGAGGAGCATATATGAAGAGTTT 361
QY 294 LysThrTyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGlu 313
DB 362 AAAACATACAGTCTTAAGTGTGACTTGGTACTTGTGCTGCACATCAGAAAGTACAGAA 421
QY 314 SerSerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp 333
DB 422 TCCAGTTTCGGGGCCAAAGAGCTGAAGAAATATGAGAGAGTGTTTTGTTCATCTCAGAC 481
```


Qy	334	PheValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThr	353
Db	482	TTCTGTTGGTACATGTTTAAAGATACAGACTCCAGTTTATGCACGAGAGATGCTTTTACT	541
Qy	354	AspSerAlaIleSerAlaLysValAsnGluGluHisLysGluLysAspLeuGluProTyr	373
Db	542	GACTTCTGCTCTCACGGCAAGGTGAAATGGTGAGCACAAAGGAGAGGACCTGGAGCCCTGG	601
Qy	374	AspAlaGlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAsn	393
Db	602	GATCGACGGGAGCTCACGGCAGCGAGGAGCTGGAG--CTGGAGAAATGATGTGCTTAAT	658
Qy	394	GlyTyrAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThr	413
Db	659	GGATGGGACCCCAATGATGATGTTTCGATATAATAGAGAAATTATGTGTGGTGTCCACA	718
Qy	414	TyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPhe	433
Db	719	TATGATAGCAGTTTATCTTCATATACCGTTCCTTTAGAAAGGGACAACCTCAGAAAGATTT	778
Qy	434	LeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyr	453
Db	779	CTTTAAACGGGAGGCAAGGGCAACACAGTTAGCAGAGAATTTGAATCCAGTGTCTCAGTAC	838
Qy	454	LysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluGluLysTyrThrAlaVal	473
Db	839	AAACCTCGTGTGCGCCTTTGAGATGATGATACCGGAGTGGGAAGAAAAAATACACAGCAGTC	898
Qy	474	GlnArgAsnSerSerGluArgGluGluGlyHisSerIleAsnThrArgGluAsnLysTyrIle	493
Db	899	CAGAGAAACTGCAGTGACCGGGAGGGGATGGCCCAACACTAGGGAATAATAATATATT	958
Qy	494	ProProGlyGlnArgAsnArgGluValIleSerTyrGlySerGlyArgGlnAsnSerPro	513
Db	959	CCTCTCTGGACAAAGAACACAGAGAAGTCTCTATCTTGGGGAGTGGGACAGAGCTCACCA	1018
Qy	514	ArgMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAsp	533
Db	1019	CGGATGGGCGCAGCCTGGGCGCAGGCTCCATGCGCTCAAGAGCTGCTCTTCACACTTCAGAT	1078
Qy	534	PheAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTyrProSer	553
Db	1079	TTCAACCCGAAACGCTGGCTCAGACCAAGAGTAGTTAATGAGGTGTTCCTCTGGCCATCG	1138
Qy	554	ProCysProSerProSerArgProProSerArgTyrGlnSerGlyProAsnSerLeu	573
Db	1139	CCTTGGCCATCTCTCTCTCGCCACCCTTCGCTACCACTCAGGTCCCAACTCTCTT	1198
Qy	574	ProProArgAlaAlaThrProThrArgPro	583
Db	1199	CCACCTCGGGCAGCAGCCCTTACACGGCGCT	1228

RESULT 10

```

US-08-981-998A-19
; Sequence 19, Application US/08981998A
; Patent No. 684431
;
; GENERAL INFORMATION:
;
; APPLICANT: PULST, STEFAN M.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
; ATAXIA-2 AND PRODUCTS RELATED THERETO
;
; NUMBER OF SEQUENCES: 19
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: MUEITING, RAASCH & GEBHARDT, P.A.
; STREET: 119 NORTH FOURTH STREET, SUITE 203
; CITY: MINNEAPOLIS
; STATE: MINNESOTA
; COUNTRY: 55401
; ZIP: 55401
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0. Version #1.30

```

```
Db 362 AAAACATACAGTCCTAAGTGTGACTTGGTACTTGTGCTGCACATGAGAAAGTACAGAA 421
Qy 314 SerSerSerGlyProlysArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp 333
Db 422 TCCAGTTCCGGGCGCAAAACGTAAGAAATAATGGAGAGTGTGTTGTTCAATGCTCAGAC 481
Qy 334 PheValValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThr 353
Db 482 TTCGTTGCTGTACAGTTTAAAGATACAGACTCCAGTTATGCACGGAGAGATGCTTTTACT 541
Qy 354 AspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrp 373
Db 542 GACTCTGCTCTCAGCCCAAGGTGAATGTTGTCACACAGAGAGAGACCTGGAGCCCTGG 601
Qy 374 AspAlaGlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAsn 393
Db 602 GATGCAGGGAGGAGCTACCGCCAGCAGGAGGCTGGAG--CTGGAGAAATGATGTGCTAAT 658
Qy 394 GlyTrpAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThr 413
Db 659 GGATGGGACCCCAATGACATGTTTCGATATATATGAAGAGAAATTTATGTTGTTGTCCACA 718
Qy 414 TyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPhe 433
Db 719 TATGATAGCAGTTTATCTTCATATACGGTTCCTTTAGAAAGGACCACTCAGAAAGATTT 778
Qy 434 LeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyr 453
Db 779 CTAAACCGGGAGGCAAGGGCAACACCATGTAGCAGAGAAATTTGAATCCAGTGTCTAGTAC 838
Qy 454 LysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluLysTyrThrAlaVal 473
Db 839 AAAGTCGTGTCCGCTTCAGAAATGATGACCGGAGTGAGGAGAAATAACACAGCAGTC 898
Qy 474 GlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIle 493
Db 899 CAGAGAAACTGCAGTACCGGGAGGGGCGATGGCCCAACACTAGGACAAATAATATATT 958
Qy 494 ProProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerPro 513
Db 959 CCTCTCTGCAAGAAACAGAGAACTCTATCTCTGGGGAAGTGGGAGACAGACTCACCA 1018
Qy 514 ArgMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAsp 533
Db 1019 CGATGGCCGAGCTGGGGCAGGCTCCATGCCGTCAAGAGCTGCTTCTCACACTTCAGAT 1078
Qy 534 PheAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTrpProSer 553
Db 1079 TTCAACCCGAAACGCTGGCTCAGACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCG 1138
Qy 554 ProCysProSerProSerArgProProProSerArgTyrGlnSerGlyProAsnSerLeu 573
Db 1139 CCTTGCCCATCTCTTCTCTGCCCCACCTTCTCGCTACCACTCAGTCCCACTCTCTT 1198
Qy 574 ProProArgAlaAlaThrProThrArgPro 583
Db 1199 CCACCTCGGGCAGCACCCCTACACGGCCT 1228
```

RESULT 11

```
US-09-949-016-16240
; Sequence 16240, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16240
; LENGTH: 151088
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(151088)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16240
```

Alignment Scores:

```
Pred. No.: 1,368-37 Length: 151088
Score: 1245.50 Matches: 354
Percent Similarity: 49.04% Conservative: 54
Best Local Similarity: 42.55% Mismatches: 176
Query Match: 17.89% Indels: 251
DB: 3 Gaps: 32
```

US-10-802-228-2 (1-1312) x US-09-949-016-16240 (1-151088)

```
Qy 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db 2163 ATGCGCTCAGCGCGCGCAGCTCTCGAGTCCCGGTGGCCACCAGTCTCCCGG-TTC 2221
Qy 21 AlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgArgSerGlyArg 40
Db 2222 GCCGAGCAGCAGGTGGCGCGGTGGCGCTCGCTCCAGCGCGCGCGCGCGCGCG 2281
Qy 41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 2282 GCGCGCGGTGGCGCGCGCGCGCGTATCCCTCCCGCGCGCGCGCGCGCGCGCC 2341
Qy 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db 2342 GCGCGCGCGCGCGCGCGCGCGCGCTCCCTCCCGCAGAGCTCGCTCCCTCCAGCTGTTTGGTAGCAAC 2401
Qy 81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyGlyPro 100
Db 2402 GGCACAGCGCGCGCGCGCGTTCGGCCCGGTCCCGCGCGCTCTTGGTCTCGCGCGCGCT 2461
Qy 101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProAlaAla 120
Db 2462 CCGCGCGCGCTTCGTGCTGCTCTCTCCCTCCCGCGCGCGCGCGCGCGCGCGCG 2521
Qy 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db 2522 CCAACCGCGCGCTCCCGCGCTCGCGCGCGCGCGCTCCCGCGCGGTTCCTCGCGCGTCTCCTTG 2581
Qy 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db 2582 GCGCGCGCGGTCTCCCGCGGTGTCGCCCGCGCGGTGGAGCGGTATGGCGCGCTCACC 2641
Qy 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db 2642 ATGTCGTGAAGCCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAACAG 2701
Qy 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
Db 2702 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2761
Qy 201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerSerValSerSer 220
Db 2762 AGCGCGCTTCTACGCTCGCGCGCGCGCGCGCTTCGCCGTCTCTCGCTCTCGCTCG 2821
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
Db 2822 TCCTCGGCGCAGGCTCCCTCTCTCGGTGTCGCGCGCGCGCGCGCGCGCGCGCGCG 2881
Qy 241 LeuGlyArg-----GlyArgAsnSerAsnLysGlyLeuPro 252
```


;; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
;; NUMBER OF SEQUENCES: 19
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
;; STREET: 119 NORTH FOURTH STREET, SUITE 203
;; CITY: MINNEAPOLIS
;; STATE: MINNESOTA
;; COUNTRY: 55401
;; ZIP: 55401
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/981,998A
;; FILING DATE: 11-May-1998
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO 97/42314
;; FILING DATE: 08-MAY-1997
;; APPLICATION NUMBER: US 08/727,084
;; FILING DATE: 08-OCT-1996
;; APPLICATION NUMBER: US 60/022,207
;; FILING DATE: 19-JUL-1996
;; APPLICATION NUMBER: US 60/017,388
;; FILING DATE: 08-MAY-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MUETING, ANN M.
;; REGISTRATION NUMBER: 33,977
;; REFERENCE/DOCKET NUMBER: 232.00010120
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 612-305-1217
;; TELEFAX: 612-305-1228
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 516 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: both
;; TOPOLOGY: both
;; MOLECULE TYPE: DNA (genomic)
;; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-08-981-998A-1

Alignment Scores:
Pred. No.: 3,328-25 Length: 516
Score: 857.00 Matches: 166
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.31% Indels: 0
DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-08-981-998A-1 (1-516)

Qy 78 GlySerAsnGlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeu 97
Db 3 GGTAGCAACGGAACCGCGCGCGCGGTTCGGCCCGGCTCCCGCGGCTCTTGGTCTC 62
Qy 98 GlyGlyProProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaPro 117
Db 63 GCGCGGCTCCCGCGCCCTTCGTCTGTCGTCCTCCCGCTCCCGCTCCCGCGCGCGCT 122
Qy 118 ProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGly 137
Db 123 CCGGCGGCGCCACCGCGGCTCCCGCTCGGCGCGCGGTGGTTCCTCCCGCGGCTCCCGC 182
Qy 138 ValSerLeuAlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGly 157
Db 183 GTCTCTTTGGCGCGCGCGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCGGCTCCCG 242
Qy 158 ProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 177

Db 243 CCCCTCACCATTGTCGTGAAGCCCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 302
Qy 178 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
Db 303 CAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 362
Qy 198 ProGlyGlySerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSer 217
Db 363 CCGCGCGCGCAGCGGCTTCTAGCGTCGCGCGCGCGCGCGCTTCGCTCGCTCGCTCG 422
Qy 218 ValSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGly 237
Db 423 GTCTCTCTGTCCTCGGCACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
Qy 238 ArgProGlyLeuGlyArg 243
Db 483 AGGCGCGCGCTGGGCAGG 500
RESULT 15
US-09-707-919A-13
; Sequence 13, Application US/09707919A
; Patent No. 6623927
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: Method for detection of human spinocerebeller ataxia 2
; FILE REFERENCE: US 443
; CURRENT APPLICATION NUMBER: US/09/707,919A
; CURRENT FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 13
; LENGTH: 459
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Description of
; OTHER INFORMATION: Artificial derived nucleotide sequence of allelic
; OTHER INFORMATION: variant of SCA2 gene
US-09-707-919A-13

Alignment Scores:
Pred. No.: 1,378-22 Length: 459
Score: 787.00 Matches: 151
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.34% Mismatches: 0
Query Match: 11.31% Indels: 0
DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-707-919A-13 (1-459)

Qy 72 SerAlaSerAspCysPheGlySerAsnGlyAsnGlyGlyGlyAlaPheArgProGlySer 91
Db 2 TCCGCTCAGACTGTTTGTGTAGCAACGCGCGCGCGCGCTTCGCGCGCGCTCC 61
Qy 92 ArgArgLeuLeuGlyLeuGlyGlyProProArgProPheValValValLeuLeuProLeu 111
Db 62 CGCGCGCTCTGTCCTCGCGCGGCTCCCGCGGCTTCGTCGCTCTCTCTCTCTCTCTCT 121
Qy 112 AlaSerProGlyAlaProProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAla 131
Db 122 GCCAGCCCGCGCGCGCTCCCGCGCGCGCAACCGCGCGCTCCCGCGCTCCCGCGCG 181
Qy 132 SerProProArgSerGlyValSerLeuAlaArgProAlaProGlyCysProArgProAla 151
Db 182 TCCCGCGCGCTTCGCGGCTCTCTCTGCGCGCGCGCTCCCGCTTCCTCTCTCTCTCTCT 241
Qy 152 CysGluProValTyrGlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGln 171
Db 242 TGCAGCGCGGTGTATGGGCCCTTCACCATGTCTCGTGAAGCCCCAGCAGCAGCAGCAG 301
Qy 172 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 191

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 23, 2005, 06:22:40 ; Search time 1990 Seconds
(without alignments)
5451.972 Million cell updates/sec

Title: US-10-802-228-2
Perfect score: 6961
Sequence: 1 MRSAAAPRSPAVATESRRF.....TAHFPYMTHPVSQAHHQQL 1312

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10802228/runat_21122005_110226_9755/app_query.fasta_1.1479
-DB=Published Applications_NA_Main -QPMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10802228@cgn_1.1 1549 @runat_21122005_110226_9755 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA_Main:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6961	100.0	4481	6	US-10-141-541-5
2	6961	100.0	4481	8	US-10-750-323-2
3	6961	100.0	4481	9	US-10-887-553A-493
4	6961	100.0	4481	9	US-10-802-228-1
5	6025.5	86.6	4225	6	US-10-141-541-3
6	2463	35.4	2006	3	US-09-925-298-15
7	2463	35.4	2006	5	US-10-102-806-15

8	1876	27.0	1257	6	US-10-141-541-1	Sequence 1, Appli
9	1876	27.0	1257	8	US-10-750-323-4	Sequence 4, Appli
10	1816.5	26.1	4674	5	US-10-171-581-346	Sequence 346, App
c 11	1042.5	15.0	1707	3	US-09-823-245A-600	Sequence 600, App
c 12	917	13.2	682	9	US-10-956-157-3976	Sequence 3976, Ap
c 13	917	13.2	682	9	US-10-956-157-9211	Sequence 9211, Ap
14	857	12.3	516	8	US-10-750-323-1	Sequence 1, Appli
15	799	11.5	485	3	US-09-918-995-28268	Sequence 28268, A
16	735	10.6	3210	10	US-11-097-143-10706	Sequence 10706, A
c 17	725	10.4	5975	10	US-11-097-143-10705	Sequence 10705, A
c 18	655	9.4	465	4	US-09-925-065A-129085	Sequence 129085,
c 19	598	8.6	493631	5	US-10-087-192-205	Sequence 205, App
c 20	587	8.4	13416	5	US-10-205-032-11	Sequence 11, Appli
c 21	587	8.4	60196	5	US-10-205-032-1	Sequence 1, Appli
c 22	568.5	8.2	64492	7	US-10-378-083-1	Sequence 1, Appli
c 23	568	8.2	18717	6	US-10-156-761-927	Sequence 927, App
c 24	568	8.2	30690	7	US-10-204-862A-1	Sequence 1, Appli
c 25	568	8.2	30690	10	US-11-005-196-1	Sequence 1, Appli
c 26	568	8.2	9025608	6	US-10-156-761-1	Sequence 1, Appli
c 27	566	8.1	11238	5	US-10-205-032-15	Sequence 15, Appli
c 28	564	8.1	88400	8	US-10-844-716-1	Sequence 1, Appli
c 29	556.5	8.0	138203	9	US-10-819-386A-1	Sequence 1, Appli
c 30	543.5	7.8	85915	7	US-10-647-196-1	Sequence 1, Appli
c 31	541.5	7.8	67167	9	US-10-937-730A-3	Sequence 3, Appli
c 32	541	7.8	29467	9	US-10-937-730A-1	Sequence 1, Appli
c 33	528	7.6	65140	7	US-10-203-295-1	Sequence 1, Appli
c 34	528	7.6	125401	7	US-10-203-295-35	Sequence 35, Appli
c 35	522.5	7.5	9025608	6	US-10-156-761-1	Sequence 1, Appli
c 36	520.5	7.5	73995	5	US-10-087-192-208	Sequence 208, App
c 37	520.5	7.5	84428	7	US-10-229-148B-1	Sequence 1, Appli
c 38	515	7.4	22017	8	US-10-211-028-12	Sequence 12, Appli
c 39	515	7.4	90597	8	US-10-211-028-1	Sequence 1, Appli
c 40	511	7.3	20922	3	US-09-942-025-14	Sequence 14, Appli
c 41	511	7.3	67311	3	US-09-942-025-1	Sequence 1, Appli
c 42	501	7.2	34094	3	US-09-963-038A-1	Sequence 1, Appli
c 43	501	7.2	34094	5	US-10-199-550-1	Sequence 1, Appli
c 44	501	7.2	34094	6	US-10-245-603A-1	Sequence 1, Appli
c 45	498.5	7.2	4725	3	US-09-940-316B-24	Sequence 24, Appli

ALIGNMENTS

RESULT 1
US-10-141-541-5
; Sequence 5, Application US/10141541
; Publication No. US20030167495A1
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M.
; TITLE OF INVENTION: SCA2 Knockout Animal and Methods of Use
; FILE REFERENCE: P-CE 5244
; CURRENT APPLICATION NUMBER: US/10/141,541
; CURRENT FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: US 60/289,231
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (163)...(4101)
US-10-141-541-5

Alignment Scores:
Pred. No.: 0
Score: 6961.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 6
Length: 4481
Matches: 1312
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-802-228-2 (1-1312) x US-10-141-541-5 (1-4481)	
Qy	1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db	163 ATGCGCTCAGCGCGCGAGCTCTCGGAGTCCGCGGTGGCCACCGAGTCTCGCGCTTC 222
Qy	21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db	223 GCCGAGCGCAGGTGGCGGTGGCGCTCGCTCCAGCGCGCGCGCGCGAGCGGGCGG 282
Qy	41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db	283 GCGCGGTGGCGCGCGCGCGGACCGTATCCCTCCGCGCGCGCTCCCGCGCGCGCCCC 342
Qy	61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db	343 GSCCCCCCTCCCTCCGCGCAGAGCTCGCTCCCTCCGCGCTCAGACTGTTTGGTAGCAAC 402
Qy	81 GlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyPro 100
Db	403 GGCACCGCGCGCGCGCGGTTCGCGCCGCGCTCCCGCGCGCTCCTTGGTCTCGCGCGGCT 462
Qy	101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAla 120
Db	463 CCCCCCGCTTCGTCTGCTCTCTCCCGCTCGCCAGCGCGCGCGCGCTCCCGCGCGG 522
Qy	121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db	523 CCACACCGCGCTCCCGCTCGCGCGCGGTGGTCCCGCGCGTCCCGCGCTCTCTTG 582
Qy	141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db	583 GCGCGCGCGCTCCCGGTGTCCCGCGCGCGGTGGTCCCGCGCGTATGGGCGCTCACC 642
Qy	161 MetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db	643 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
Qy	181 GlnGlnGlnGlnGlnGlnProProProAlaAlaAlaAsnValArgLysProGlyGly 200
Db	703 CAGCAGCAGCAGCAGCAGCAGCGCGCGCGCGCGGTGCCATGTCCGCAAGCCCGCGC 762
Qy	201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerValSerSer 220
Db	763 AGCGGCTTCTAGCTGCGCGCGCGCGCGCTTCGCGGTCTCGCTCTCGCTCTCTCG 822
Qy	221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGlyArgProGly 240
Db	823 TCCTCGGCCACGGCTCCCTTCGTGGTTCGCGCGACCTCCGCGCGCGGAGGCGCGCGC 882
Qy	241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
Db	883 CTGGCAGAGGTGGAACAGTAACAAAGAGCTGCCTCAGTCTACGATTTCTTTGATGGA 942
Qy	261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValGlySerLysCysGlu 280
Db	943 ATCTATGAAAATATGAGGATGTTTCATATACATCATCAGTTGTGGCTCCAAATGTGAA 1002
Qy	281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db	1003 GTACAAAGTGAAAATGAGGATATATGAAGAGGATTTTAAACCTTACAGTCGGAAGTGT 1062
Qy	301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
Db	1063 GATTTGGTACTTGTATGCCGCACATGAGAAAAGTACAGAAATCCAGTTCGCGGCGGAAACGT 1122
Qy	321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
Db	1123 GAAGAAAATAATGGAGAGTATTTTGTTCAAATGTTTCAGACTTTGTGTGTACAGTTTAAA 1182
Qy	341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db	1183 GATATGGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTCTAAA 1242
Qy	361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAla 380
Db	1243 GTGAATGCGCAACACAAAGAGAAGACTTGGAGCCCTGGGATCGAGGTGAACCTCACAGCC 1302
Qy	381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
Db	1303 AATGAGAACTTGAGGCTTTGGAAAATGACCGTATCTAATGATGGATGCCAATCATATG 1362
Qy	401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
Db	1363 TTTTCATATAATGAAGAAAATTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
Qy	421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db	1423 TATACAGTGCCTTTAGAAAAGAGATAACTCAGAAGAAATTTTAAACCGGAAGCAAGGCA 1482
Qy	441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db	1483 AACCAAGTTAGCAGAAGAAATTTGAGTCAAGTCCCGAGTACAAAGCTCGAGTGGCCCTGGA 1542
Qy	461 AsnAspAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db	1543 AATGATGATAGAGTGAGGAAGAAAATACACAGCAGTTCAGAGAAATTCAGGTGAACGT 1602
Qy	481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
Db	1603 GAGGGGCACAGCATAAACACTAGGAAAATAAATAATATATCTCTCTGGACAAAGAAATAGA 1662
Qy	501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db	1663 GAAGTCATATCTCTGGGAAGTGGGAGACAGAAATTCACCGGTATCGGCGCAGCTGGATCG 1722
Qy	521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db	1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGAAATCTGGTTCA 1782
Qy	541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560
Db	1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCTTGCCTCTCTCTCTCTCT 1842
Qy	561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db	1843 CGCCACCTTCTCGCTACCACTAGTCCCAACTCTCTTCACCTCGGCGCACCCCT 1902
Qy	581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db	1903 ACACGCGCGCTCCAGGCCCTCCGGGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
Qy	601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db	1963 CATGTTCTCCAGCTCTCTGTCTCTACTATGCTTAAACGATGTCTTCAGAAAGGCGCTCCA 2022
Qy	621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Db	2023 AGGATGTCCCAAGGCGCAGCAGCATCTCGAAATCACAGAGTTTCTGCTGGGAGGGGT 2082
Qy	641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
Db	2083 TCCATATCCAGTGGCCCTAGAAATTTGTATCCCAACCCACCCAGTGGTGAAGCAGTACTCT 2142
Qy	661 ProValAlaArgThrSerProSerGlyGlyThrTyrPheSerSerValSerGlyValPro 680
Db	2143 CCAGTAGCAAGGACAGTCCCTCGGGGGAACGTGGTTCATCAGTGGTTCAGTGGGTTCCA 2202
Qy	681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db	2203 AGATTATCCCTTAAACCTCATAGACCAGGCTCTCCACAGACAGACAGTATTGGAAATACC 2262
Qy	701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
Db	2263 CCCAGTGGGCCAGTCTCTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGCC 2322

QY 721 MetProIleProAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
DB 2323 ATGCGCTATTCCAGCTGCATCTCTAGCGCTGTAGTCTGCGATCGAACAGAGCTGTATCC 2382
QY 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
DB 2383 CTTCTAGTGGGCTAAGATTCAGGCTTCAGATCAGAGGCGAGNACTCTCTCGCAGGG 2442
QY 761 AsnLysGluAsnLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
DB 2443 AATAAGAAAAATATAACCCAAATGAACATCACCTAGCTTCTCAAAGCTGAAACAAA 2502
QY 781 GlyLysSerProValValSerGluHisArgLysGlnLysAspLysLysPheLys 800
DB 2503 GGTATATCATCCAGTGTCTGTAACATAGAAACAGATTGATTTAAGAAATTTAAG 2562
QY 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
DB 2563 ATGATTTTAGGTTACAGCCAGTCTCTCTGATCTATGATCAACTACTNAACAAA 2622
QY 821 AsnArgGluGlyGluLysSerArgAspLeuLysLysLysLysLysLysLysLys 840
DB 2623 AATAGAGAGGAGAAAAATCAAGAGATTGATCAAGAGCAAAATTCGAA CCAAGTGTAAAG 2682
QY 841 AsnSerPheLysGlnAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
DB 2683 GATTTCTTTTCAATGAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCGGAATAGC 2742
QY 861 ProSerLysSerProSerLysLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
DB 2743 CCAGCATTTCCCTTCAATCTAGTAAACAGGAGCAGCAAGAGGGGAGCTGAGGTCACT 2802
QY 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnLysAspLysLysGlu 900
DB 2803 TCCCAAGGGTTCAGACTTCCAGCCAGCATGTAAACAGAGAAACGATTAAGGAAGAG 2862
QY 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysLysLys 920
DB 2863 AAGAAAGAGCGCAGCTGAGCAAGTAGGAAATCAACATTGAATCCCAATGCAAGGAGTTC 2922
QY 921 AsnProArgSerPheSerGlnProLysProSerThrProThrSerProArgProGln 940
DB 2923 AACCCAGTTCCTTCTCAGCCAAAGCCTTCTACTACCCCAACTTCACTCGGCTCAA 2982
QY 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValThrGlnPro 960
DB 2983 GCACAACTAGCCCATCTATGTGTGGTTCATCAACAGCCAACTCCAGTTTATCTCAGCCT 3042
QY 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
DB 3043 GTTTGTTTGTGCACAAATATGATGTATCCAGTCCCGAGTGGCCAGCGCGTCAACCTTTA 3102
QY 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
DB 3103 TACCCAACTATGATGAGCCCATGCCAGTAGTAATCAAGCCAAAGACATATAGAGCAGTACCA 3162
QY 1001 AsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMetHisProAlaSer 1020
DB 3163 AATATGCCCAACAGCGGCAAGCAGCAGCATCATCAGATGCTGATGATGCCAGCGTCA 3222
QY 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
DB 3223 GCAGCGGCCCAACCGATTGAGCCAGCCACCCAGCTTACTCCAGCGCAATATGTGCTAC 3282
QY 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
DB 3283 AGTCCTCAGAGTTCCCAATCAGCCCTTGTTCAGCATGTGCCAATATCATGTCTCAG 3342
QY 1061 HisProHisValTyrSerProValLysGlnGlnAlaArgMetMetAlaProProThr 1080
DB 3343 CATCTCATGTCTATAGTCTGTATATACAGGTAATGCTAGATGATGGCAGCCACCAACA 3402
QY 1081 HisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100

DB 3403 CACGCCACCGCTGGTTTAGTATCTTCTTACGAACTCAGTACGGGCTCATGAGCAGACG 3462
QY 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
DB 3463 CATGCGATGTATGATGCTGCCAAATTTACCATACAAGGAGACAAGCCCTTCTTCTTCTAC 3522
QY 1121 PheAlaLysSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
DB 3523 TTTGCGCATTTCCAGCGGCTCCCTGCTCAGCAGTATGCGCACCTTAAGCTTACCCTGTCAC 3582
QY 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
DB 3583 CCACATATCTCCACACCTCAGCCTTACGCTACCCCACTGGACAGCAGCAAGCAACAT 3642
QY 1161 GlyLysSerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
DB 3643 GGTGGAAGTCACTCTGCAACCCAGTCTGTTCAGCACCATCAGCACCGCCGCCAGGCT 3702
QY 1181 LeuHisLeuAlaSerProGlnGlnGlnSerAlaLysThrHisAlaGlyLeuAlaProThr 1200
DB 3703 CTCCATCTGCGCAGTCCACAGCAGCAGTACGCAATTTACCACGCGGGCTTGGCCAACT 3762
QY 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
DB 3763 CCACCTTCATGACACCTGCTCCACACGAGTCCACAGCAAGTATTTCCAGCAGCA 3822
QY 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
DB 3823 CAACAGACTGTCTTTACGATCCATCTCTCAGTTCAGCCGCGGTATACCAACCCACCC 3882
QY 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
DB 3883 CACATGGCCACGCTACTCAGGCTCATGTACAGTCCAGGAATGTTCTCTCATTCACCT 3942
QY 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
DB 3943 GCCCATGCCCAATGATGCTAATGACACACAGCCACCCCGCGGTCCCGAGGCCGCTC 4002
QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
DB 4003 GCTCAAGTGCATACAGGCCATTCAGTCTCGACAACAGCGCATTTCCCTATATGACG 4062
QY 1301 HisProSerValGlnAlaHisHisGlnGlnLeu 1312
DB 4063 CACCTCTAGTACAAGCCCAACCAACAGCAGTTC 4098

RESULT 2

US-10-750-323-2
; Sequence 2, Application US/10750323
; Publication No. US20050032083A1
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBELLAR
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Mueeting, Raasch & Gebhardt, P.A.
; STREET: 119 North Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 30-Dec-2003
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/727,084
FILING DATE: 08-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Muetting, Ann M.
REGISTRATION NUMBER: 33,977
REFERENCE/DOCKET NUMBER: 232.00010101
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612/305-1220
TELEFAX: 612/305-1228
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4481 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 163..4101
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-750-323-2

Alignment Scores:
Pred. No.: 0 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-802-228-2 (1-1312) x US-10-750-323-2 (1-4481)

1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
163 ATGCGCTCAGCGCGCGAGCTCTCGGAGTCCCGGCTGGCCACCGAGCTCTCGCGCTTC 222
21 AlaAlaAlaArgTirProGlyTirArgSerLeuGlnArgProAlaArgSerGlyArg 40
223 GCCGAGCAGCGTGGCGCGGCTCGCTCCAGCGCGCGCGCGCGCGCGCGCGCG 282
41 GlyGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
283 GCGCGCGTGGCGCGCGCGCGCGCGTATCCCTCCCGCGCGCGCGCGCGCGCGCG 342
61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
343 GCGCGCGCTCCCTCCCGCGAGAGCTCGCTCCCTCCGCTCAGACTGTCTTGGTAGCAAC 402
81 GlyAsnGlyGlyAlaAlaPheArgProGlySerArgArgLeuLeuGlyGlyGlyPro 100
403 GCGAAGCGCGCGCGCGCTTCCGCGCGCGCTCCGCGCGCTCCTTGGTCTCGCGCGGCT 462
101 ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAlaAla 120
463 CCGCGCGCGCTTCGTCTGTCTTCTCCCGCTCCGCGCGCGCGCGCGCGCGCGCG 522
121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
523 CCNACCGCGCGCTCCCGCTCGCGCGCGCTGGCTCCCGCGCGGTTCCGGCGTCTCCCTG 582
141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
583 GCGCGCGCGCTCCCGCTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 642
161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
643 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702
181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 200
703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762
201 SerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSerValSerSer 220

Db 763 AGCGCCTTCTAGCGTCCCGCGCGCGCTTCCGCGCTCTCGCTCTCGCTCTCTCTCG 822
Qy 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyGlyArgProGly 240
Db 823 TCCTCGCGCACGGCTCCCTCTCGTGGTCCGCGGACCTCCGCGGCGGAGCGCGCGC 882
Qy 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrLysSerPheAspGly 260
Db 883 CTGGCAGAGGTCGAAACAGTAACAAAGGACTGCCTCAGTCTAGCATTTCTTTTGTATGA 942
Qy 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
Db 943 ATCTATGCAAAATATGAGGATGGTTTATATATGAGGAGGTTTATAAACTTACAGTCC 1002
Qy 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db 1003 GTACAAGTGAAATATGAGGATGATATATGAGGAGGTTTATAAACTTACAGTCC 1062
Qy 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
Db 1063 GATTTGGTACTTGTATGCGCACATGAGAAAGTAGTACAGATCCAGTTCGGGCGCGAAACGT 1122
Qy 321 GluGluLeuMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
Db 1123 GAAGAAATATATGAGGAGTATTTTCTCAAAATGTTTACAGACTTTGTGTGACAGTTTAAA 1182
Qy 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db 1183 GATATGGAATCCAGTTATGCAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGTCTAAA 1242
Qy 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTirPheLysGluLeuThrAla 380
Db 1243 GTGAATGCGCAACACAAAGAGAGACTGAGGCGCTGGAGCTGGAGTGAACACTCACAGCC 1302
Qy 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTirPheAspProAsnAspMet 400
Db 1303 AATGAGGAATCTGAGGCTTTGGAAATACAGTATCTAATGAGTGGATCCCAATGATATG 1362
Qy 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
Db 1363 TTTTCATATATGAGAAATAATTTATGGTGTAGTGTCTAGTATGATAGCAGTTTATCTTCG 1422
Qy 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db 1423 TATACAGTGCCTTTAGAAAGAGATAACTCAGAACAAATTTTAAACGGGAAGCAAGGCA 1482
Qy 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db 1483 AACCAAGTTAGCAGAAATAATTTAGTCAAGTCCCAAGTACAAAGCTCGAGTGGCGCTGAA 1542
Qy 461 AsnAspArgSerGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db 1543 AATGATGATAGGAGTGAGGAAGAAATAATACACAGCAGTTTCAGAAATTTCCAGTCAACGT 1602
Qy 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProGlyGlnArgAsnArg 500
Db 1603 GAGGGGCACAGCATTAACACTAGGGAAATAAATAATATATCTCTCGGACAAAGAAATAGA 1662
Qy 501 GluValIleSerTirPheGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db 1663 GAAGTCATATCTCGGGAAGTGGGAGAGCAGAAATTCACCGCTATGGGCGAGCTGGATCG 1722
Qy 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db 1723 GGCTCCATGCCATCAAGATCCACTTTCACACTTTCAGATTTCAACCCCGAAATCTTGTTCA 1782
Qy 541 AspGlnArgValValAsnGlyGlyValProTirPheProSerProCysProSerProSerSer 560
Db 1783 GACCAAGAGTAGTTAATGGAGGTGTTCCCTGGCCATCGCTTCCCATCTCTCTCTCTCT 1842
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580

1843 CGCCACCTTCTCGCTACAGTCAGTCCCAACTCTCTTCCACCTCGGCGACGCCACCTT 1902
581 ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla 600
1903 ACAGGGCCCTCCAGGCCCTCCAGGCCCTCCAGGCCCTCCAGGCCCTCCAGGCCCTCCAG 1962
601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
1963 CATGGTCTCCAGCTCTCTCTACTATGCTTAAACGCGATGCTTTCAAGAGGGCTCCA 2022
621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
2023 AGGATGTCCTCCAAAGGCCCGACGACATCTCTGNAATCAGAGATTCTCTCTGGAGGGT 2082
641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
2083 TCCATATCCAGTGGCTAGAAATTTGATCCCAACCCAGCCAGTGAAGCAGCTACTCTCT 2142
661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680
2143 CCAGTAGCAAGGACCACTCTCTCGGGGGAAACGTGTCATCATGTCAGTGGGGTTC 2202
681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
2203 AGATTATCCCTTAAACTCATAGACCAGGTCTCCAGACAGACAGTATTTGGAATATACC 2262
701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
2263 CCCAGTGGGCCAGTCTCTGCTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGCC 2322
721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
2323 ATGCCTATTCCAGCTGCATCTCTACGCCCTGTAGTCTCTGATCGAACAGAGCTGTACC 2382
741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
2383 CCTTCTAGTGAGGCTAAAGATTCCAGGCTTCAAGATCAGAGGCAGAACTCTCTCTGAGGG 2442
761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
2443 AATAAGAAATATTAACCCATGAACATCATCACTAGCTTCTCAAAAGCTGAACACAA 2502
781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLysLysPheLys 800
2503 GGTATATCACAGTGTCTTCTGAACATAGAAACACAGATTGATGATTTAAGAAATTAAG 2562
801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
2563 AATGATTTTAGGTACAGCCCAAGTCTTACTTCTGAATCTATGGATCAACTACTTAAACAA 2622
821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
2623 AATAGAGAGGAGAAATCAAGAGATTGATCAAGACAAATTTGAACCAAGTGCTAAG 2682
841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
2683 GATTCTTTCATGAAATAGCAGCAGCAACTGTACCAGTGGCAGCAGCAAGCCGAAATAGC 2742
861 ProSerIleSerProSerIleLeuSerAsnThrCluHisArgGlyProGluValThr 880
2743 CCCAGCATTTCCCTCTCAATCTTAGTAACAGGAGCACAGAGGGGACCTTAGGTCACT 2802
881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900
2803 TCCCAAGGGTTCAGACTTCCAGCCAGCATGTAAACAAAGAGAAAGCATAGGAGAG 2862
901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
2863 AAGAAAGACGCGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAAAGGAGTTC 2922
921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
2923 AACCCAGTTCCTTCTCTAGCCCAAGCCCTTCTACTACCCCAACTTCACTCGGCCCTCAA 2982

941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValThrThrGlnPro 960
2983 GCACAACCTAGCCCATCTATGGTGGGTATCAACAGCAACTCCAGTTTATTACTCAGCCT 3042
961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
3043 GTTTGTTTGGCACCAAAATATGATGATCCAGTCCAGTGAGCCAGGGCGTCAACCTTTA 3102
981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
3103 TACCCAATACCTATGACGCCCATGTCAGTGAATCAAGCAACACATATAGAGCAGTACCA 3162
1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
3163 AATATGCCCCAACACGCGCAAGCAGCATCATCAGAGTGCCATGATGACCCAGCGTCA 3222
1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
3223 GCAGGGGGCCACCATGTCAGCCACCCACAGCTTACTCCAGCAATATGTTGCTAC 3282
1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
3283 AGTCCTCAGCAGTCTCCAAATCAGCCCTTGTTCAGCATGTGCCATATCAGTCTCAG 3342
1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
3343 CATCTCATGTCTATAGTCTCTAATACAGGGTAATGCTAGAAATGATGACCAACCA 3402
1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
3403 CAGGCCAGCCTGGTTTAGTATCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462
1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
3463 CATGCGATGATGATGTCATGCCAAATTACCATAACAAGAGACAAGCCCTTCTTCTAC 3522
1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
3523 TTTGCGATTCCACGGGCTCCCTTGCTCAGCAGTATGCCACCCCTAACGCTACCTGTCAC 3582
1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
3583 CCATATCTCCACACCCTCAGCCTTCCAGTACCTACCCCACTGGACAGCAGCAAGCCAA 3642
1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
3643 GGTGGAAGTCATCTCTGCCACCCAGTCTGTTTCAGCACCATCAGCACCGCCCGCCAG 3702
1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
3703 CTCATCTGGCCAGTCACAGCAGCAGTCCAGCCATTTACCACCGGGGCTTGGCGCAACT 3762
1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
3763 CCACCTTCCATGACACTGCTCCAACAGCAGTCGCGCCACAGAAATAGTTTCCAGCAGCA 3822
1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
3823 CAACAGACTGTCTTTACGATCCATCTCTCAGCTTCCAGCCGGGTATACCAACCCACCC 3882
1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
3883 CACATGGCCCACTAGCTCAGGCTCATGATCAGTCAGGAATGGTTCTCTCTCATTCACAT 3942
1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyProGlnAlaAlaLeu 1280
3943 GCCCATGGCGCAATGATGCTAATGACACACGCCCGCGGCTCCCCAGGCGCCCTC 4002
1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
4003 GCTCAAGTGCATACAGCCCATTCAGTCTCCACAACAGCGCATTTCCCTATATGACG 4062


```
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
Db |||||
4063 CACCCTTCAGTACAAGCCCAACCAACAGCAGTTG 4098

RESULT 3
US-10-887-553A-493
; Sequence 493, Application US/10887553A
; Publication No. US20050085436A1
; GENERAL INFORMATION:
; APPLICANT: Garza, Dan
; APPLICANT: Li, Hao
; TITLE OF INVENTION: Method to treat conditions associated
; with insulin signalling dysregulation
; FILE REFERENCE: 4-3262
; CURRENT APPLICATION NUMBER: US/10/887, 553A
; CURRENT FILING DATE: 2004-07-08
; PRIOR APPLICATION NUMBER: 60/485,883
; PRIOR FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 1208
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: human
US-10-887-553A-493

Alignment Scores:
Pred. No.: 0 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-802-228-2 (1-1312) x US-10-887-553A-493 (1-4481)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
Db |||||
163 ATGCGCTCAGCGCGCGCAGCTCTCGGAGTCCCGCGGTGGCCACCGAGTCTCGCGCTTC 222

QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgSerGlyArg 40
Db |||||
223 GCCGACGACAGTGGCCCGCGGTGGCGCTCGCTCCAGCGCGCGCGCGGAGCGGGCGG 282

QY 41 GlyGlyGlyValAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db |||||
283 GCGCGCGTGGCGCGCGCGCGGACCGTATCCTCCGCGCGCCCTCCCGCGCGCGCGCC 342

QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
Db |||||
343 GCGCGCGCTCCCTCCCGCAGAGCTCGCTCCTCCGCTCCAGCTGTGTTGGTAGCAAC 402

QY 81 GlyAsnGlyGlyValAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyPro 100
Db |||||
403 GCAACGCGCGCGCGCGCGGTTCGCGCGCGCTCCCGCGCGCTCTTGGTCTCGCGCGGCT 462

QY 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyValAlaProProAlaAla 120
Db |||||
463 CCGCGCGCTTCGTGCTGCTCTCTCCCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCG 522

QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
Db |||||
523 CCAACCGCGCTCCCGCTCGCGCGCGGTGCGTCCCGCGCGCGGTTCGCGCGCTCTCTCT 582

QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
Db |||||
583 GCGCGCGCGCTCCCGCTGCTCCCGCGCGGTGCGAGCGGTATGGCGCGCTCTCTCT 642

QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db |||||
643 ATGTCGTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 702

QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 900
Db |||||
901 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 960
```

```
Db |||||
703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 762

QY 201 SerGlyLeuLeuAlaAlaProSerProAlaAlaProSerProSerSerSerValSerSer 220
Db |||||
763 AGCGGCTTCTAGCGTCCCGCGCGCGCGCTTCCGCGCTCTCTCTCTCTCTCTCTCTCT 822

QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
Db |||||
823 TCCTCGGCGCAGCGCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 882

QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
Db |||||
883 CTGGCAGAGGTGCAAAACAGTAACAAAGAGCTGCTCAGTCTACGATTTCTTTTATGGA 942

QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
Db |||||
943 ATCTATGCAATATAGAGATGTTTATATATCTTACATCAGTTGTGGCTCCAAATGTGAA 1002

QY 281 ValGlnValLysAsnGlyGlyIleTyrGlyValPheLysThrTyrSerProLysCys 300
Db |||||
1003 GTACAAGTGAAAAATGGAGGTATATATGAGGAGTTTTTTTAAACTTTACAGTCCGAAGT 1062

QY 301 AspLeuValLeuAspAlaAlaHisGlyLysSerThrGluSerSerSerGlyProLysArg 320
Db |||||
1063 GATTTGGTACTTGTATGCGCGCATGAGAAAAAGTACAGAAATCCAGTTCGGGCGCAAGCT 1122

QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
Db |||||
1123 GAAGAAATATATGGAGATTTTGTTCAAATGTTTCAGACTTTGTTGGTACAGTTTATAA 1182

QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
Db |||||
1183 GATATGGACTCCAGTTATGCAAAAGAGATCTTTTACTGACTCTGCTATCAGTCTCTAAA 1242

QY 361 ValAsnGlyGluHisLysGlyLysAspLeuGluProTrpAspAlaGlyGluLeuThrAla 380
Db |||||
1243 GTGAATGGCAACACAAAGAGAGAGCCCTGGAGCCCTGGGATGCAGGTGAATCTACAGCC 1302

QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
Db |||||
1303 AATGAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGATGGGATCCCAATGATATG 1362

QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
Db |||||
1363 TTTTCGATATATGAAGAAAAATTTATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422

QY 421 TyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
Db |||||
1423 TATACAGTGCCTTAGAAAGAGATAACTCAGAGAAATTTTAAACCGGGAAGCAAGGCA 1482

QY 441 AsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
Db |||||
1483 AACAGTTAGCAGAGAAATTTGAGTCAAGTCCCGAGTCAAAAGCTCGAGTGGCCCTCGAA 1542

QY 461 AsnAspAspArgSerGluGluGlyLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
Db |||||
1543 AATGATGATAGAGTGGAGAGAAATAACACAGCAGTTTCAGAGAAATTTCCAGTCAAGCT 1602

QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
Db |||||
1603 GAGGGCAGCAGCATAAAACACTAGGAAAAATAATATATCTCTCTGGCAAAAGAAATAGA 1662

QY 501 GluValIleSerTyrGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
Db |||||
1663 GAAGTCTATCTCTGGGAAAGTGGAGACAGAAATTCACCGGTATGGGCGACCGCTGGATCG 1722

QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
Db |||||
1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCCAATTTCTGGTTCA 1782

QY 541 AspGlnArgValValAsnGlyGlyValProTrpProSerProCysProSerProSerSer 560
Db |||||
```

1783 GACCAAGAGTAGTTAATGGAGGTGTTCTCCCTGGCCATCCGCTTGCCCATCTCCTTCCTCT 1842
Db
561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Qy
1843 CGCCCACTTCTCGCTACCACTGAGTCCCACTCTCTCCACCTCGGGCAGCCACCT 1902
Db
581 ThrArgProProSerArgProProSerArgProProSerArgProProSerHisProSerAla 600
Qy
1903 ACACGGCCGCTCCAGGCGCCCTCGCGGCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
Db
601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Qy
1963 CATGGTCTCCAGCTCTGCTCTCTACTATGCCCTAAACGATGCTCTCAGAGGGCTCCA 2022
Db
621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Qy
2023 AGGATGTCCTCCAAAGCCAGGACATCCTCGAATCACAGAGTTTCTGCTGGAGGGGT 2082
Db
641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerSerGluAlaAlaThrPro 660
Qy
2083 TCCATATCCAGTGGCTAGAAATTTGATCCCAACACCCAGTCAAGCAGTACTCTCT 2142
Db
661 ProValAlaArgThrSerProSerGlyGlyThrTrpSerSerValValSerGlyValPro 680
Qy
2143 CCAGTAGCAAGGACCACTGCTCGGGGGAACGTGTCATCAGTGTCTCAGTGGGTGCCA 2202
Db
681 ArgLeuSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Qy
2203 AGATTATCCCTTAAACTCATAGACCAGGTCTCCAGACAGACAGTATGGAAATACC 2262
Db
701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAla 720
Qy
2263 CCAGTGGGCGAGTTCTGCTCTCCCAAGCTGGTATTATCCAACTGAACTGTGGC 2322
Db
721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerAsnArgAlaValThr 740
Qy
2323 ATGCCATTCCAGCTGCATCTCTAGCGCTGTAGTCCGTGATCGAACAGAGCTGTACC 2382
Db
741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
Qy
2383 CCTCTAGTGAGGCTAAGATTCAGGCTTCAGATCAGAGGCGAGAACTCTCTCGAGGG 2442
Db
761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
Qy
2443 AATAAGAAATATTAAACCAATGAACATCACTACCTAGCTTCTCAAAAGCTGAAAAACAA 2502
Db
781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysPheLys 800
Qy
2503 GGTATATCACCACTGTTCTGAACATAGAAACAGATTGATGATTTAAAGAAATTTAAG 2562
Db
801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Qy
2563 AATGATTTTAGTTACGCCAGGTTCTACTTCTGAATCTATGATCACTACTATAACAA 2622
Db
821 AsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
Qy
2623 AATAGAGAGGAGAAAAATCAAGAGATTTGATCAAGACAAAAATTTGAACCAAGTCTAAG 2682
Db
841 AspSerPheIleGluAsnSerSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Qy
2683 GATTCTTTCAATTGAAATAGCAGCAGCACTGTACCAGTGGCAGCAGCAAGCCGAAATAGC 2742
Db
861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Qy
2743 CCAGAGATTTCCCTCTCAATCTTAGTAAACAGGAGCAGAGGAGGACCTGAGGTCACT 2802
Db
881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspAspLysGluGlu 900
Qy
2803 TCCCAAGGGTTCAGACTTCCAGCCCGCAGCATGTAAACAGAGAAAGCAGTAAGGAAGAG 2862
Db
901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Qy
2863 AAGAAAGACGAGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAAAGGAGTTC 2922
Db

921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Qy
2923 AACCCACGCTTCTTCTCAGCCAAAGGCTTCTACTACCCCAACTTCACCTCGGCTCAA 2982
Db
941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Qy
2983 GCACAACCTTAGCCCATCTATGGTGGGTATCAACAGCCCACTCCAGTTTATCTCAGCCT 3042
Db
961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Qy
3043 GTTGTGTTTGACCAAAATATGATGATCCAGTCCAGTCCAGCGGCGTGCAACCTTTA 3102
Db
981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Qy
3103 TACCCAAATACCTATGACGCCCATGCGCAGTGAATCAAGCCAAGACATATAGACAGTACCA 3162
Db
1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Qy
3163 AATATGCCCCAACAGCGGCAAGCAGCATCATCAGATGCCATGATGACCCAGGCTCA 3222
Db
1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Qy
3223 GCAGCGGGCCCCACCCAGTTCACCCACCCAGCTTACTCCACCAATATGTTGCCCTAC 3282
Db
1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Qy
3283 AGTCTCTCAGCAGTTCCTCAAAATCAGCCCTTGTTCAGCATGTGCCATTTATCAGTCTCAG 3342
Db
1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Qy
3343 CATCTCTCATGTCTATAGTCTCTGTAATACAGGGTAATGTAGATGATGACACCCACCA 3402
Db
1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThr 1100
Qy
3403 CAGGCCACGCTGTTTAGTATCTTCTTCAGCAACTCAGTACGGGCTCATGAGCAGACG 3462
Db
1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Qy
3463 CATGCGATGATGTCATGTCCTCAAAATTACCATACAACAGGAGACAAGCCCTTCTTCTAC 3522
Db
1121 PheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Qy
3523 TTTGCGATTTCCACGGGCTCCCTTCTCAGCAGTATGCGCACCTAACGCTACCTCTGCAC 3582
Db
1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHis 1160
Qy
3583 CCACATCTCCACACCTCAGCTTCAGCTACCCCTCAGCAGCAGCAGCAACCAACAT 3642
Db
1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
Qy
3643 GGTGGAAGTCATCTGTCACCCAGTCTGTTTCAGCACCATCAGCACCGCGCCGCT 3702
Db
1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Qy
3703 CTCCATCTGGCAGTCCACAGCAGCAGTCCAGCCATTTACCGCGGGCTTGGCCAACT 3762
Db
1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Qy
3763 CCACCTCTCATGACCTGCTTACGATCCATCTCTCAGCTTCAGCGCGGTATACCAACCCACCC 3822
Db
1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Qy
3823 CAACAGCTGTCTTTACGATCCATCTCTCAGCTTCAGCGCGGTATACCAACCCACCC 3882
Db
1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Qy
3883 CACATGGCCACGCTACCTCAGGCTCATGATCAGTCAAGAAAGTGTCTCTCTCATCAACT 3942
Db
1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Qy
3943 GCCCATGCGCAATGATGCTTAATGACACACAGCCACCCCGCGGTCCCCAGCGCGCTC 4002
Db

QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
DB 4003 GCTCAAGTGCACTACAGCCATTCAGTCTCGACAACAGCGCAATTCCTCCATATGACG 4062
QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
DB 4063 CACCCTTCAGTACAAGCCACCACCAACAGCAGTTG 4098
RESULT 4
US-10-802-228-1
; Sequence 1, Application US/10802228
; Publication No. US20050209178A1
; GENERAL INFORMATION:
; APPLICANT: Pulsat, Stefan M
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF OBESITY
; FILE REFERENCE: 825466-100151
; CURRENT APPLICATION NUMBER: US/10/802, 228
; CURRENT FILING DATE: 2004-03-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 4481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-802-228-1
Alignment Scores:
Pred. No.: 0 Length: 4481
Score: 6961.00 Matches: 1312
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-802-228-2 (1-1312) x US-10-802-228-1 (1-4481)
QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPhe 20
DB 163 ATGCGCTCAGCGCGCGAGCTCTCGAGTCCGCGGTGGCCACCGAGTCTCGCGCTTC 222
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgArgSerGlyArg 40
DB 223 GCCGACGACAGTGCCCGGGTGGCGCTCGCTCCAGCGCGCGCGCGCGAGCGGGCGG 282
QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
DB 283 GCGCGGGTGGCGCGCGCGCGGACCGTATCCCTCGCGCGCGCGCGCGCGCGCGCG 342
QY 61 GlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySerAsn 80
DB 343 GCGCGCGCTCCCTCCCGGACAGCTCGGCTCCCTCCGCTCAGACTGTTTGGTAGCAAC 402
QY 81 GlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeuGlyGlyPro 100
DB 403 GCAACGCGCGCGCGCGCTTCGCGCGCGCTCCGCGCGCTCCCTGGTCTCGCGCGCGCT 462
QY 101 ProArgProPheValValLeuLeuProLeuAlaSerProGlyValaProProAlaAla 120
DB 463 CCGCGCGCTTCGTCGTCCTTCCTCCCTTCGCGCGCGCGCGCGCGCGCGCGCGCG 522
QY 121 ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeu 140
DB 523 CCAACCGCGCTCCCGCTCGCGCGCGTGGTCCCGCGCGCTCCCGCGCGCTCCCTTCG 582
QY 141 AlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyProLeuThr 160
DB 583 GCGCGCGCGCTCCCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCGCGCGCTTCGCG 642
QY 161 MetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180
DB 643 ATGTCGCTGAAGCCCGACGACGACGACGACGACGACGACGACGACGACGACGACG 702
QY 181 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 900

DB 703 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 762
QY 201 SerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerSerSerSerSerSerSer 220
DB 763 AGCGGCTTCTAGGCTGCGCGCGCGCGCGCGCTTCGCGCTTCGCTTCGCTTCCTCG 822
QY 221 SerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgProGly 240
DB 823 TCCTCGGCGACGGCTCCCTCCCTCGTGGTGGCGGACCTCCGCGCGCGGAGCGCGCG 882
QY 241 LeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheAspGly 260
DB 883 CTGGGCGAGGCTCGAAACAGTAACAAAGGACTGCCTCAGCTACGATTTCTTTTGTGGA 942
QY 261 IleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysCysGlu 280
DB 943 ATCTATGCAATATAGGATGGTTTCATATCTACATCAGTTGTTGGCTCCAAATGTGAA 1002
QY 281 ValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
DB 1003 GTACAAGTGAATAATGGAGGTATATATGAAGGAGCTTTTAAACTTACAGTCCGAAGT 1062
QY 301 AspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysArg 320
DB 1063 GATTTGGTACTTGTATGCGGCACATGAGAAAGTACAGAACTCCAGTTCGGGCGCGAAAC 1122
QY 321 GluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
DB 1123 GAAGAAATATAGGAGAGTATTTTGTTCAAATGTTTCAGACTTTGTTGGTACAGTTTAAA 1182
QY 341 AspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLys 360
DB 1183 GATATGACTCCAGTTATGCAAAAAGAGATGCTTTTACTGACTCTGCTATCAGTGCTAAA 1242
QY 361 ValAsnGlyGluHisLysGluLysAspLeuGluProTyrPheAlaGlyGluLeuThrAla 380
DB 1243 GTGAATGCGAACACAAAGAGAGAGGAGCTCGAGCGCTGGGATGAGGTTGAATCAGAGCC 1302
QY 381 AsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspProAsnAspMet 400
DB 1303 AATGAGGAACCTTGAGGCTTTGGAAAAATGACGTATCTAATGGATGGGATCCCAATGATG 1362
QY 401 PheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSer 420
DB 1363 TTTTCGATATATGAAGAAAAATATGGTGTAGTGTCTACGTATGATAGCAGTTTATCTTCG 1422
QY 421 TyrThrValProLeuGluUArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla 440
DB 1423 TATACAGTGCCTTTAGAAAGAGATAACTCAGAAGAAATTTTAAACCGGAGCAAGGCA 1482
QY 441 AsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGlu 460
DB 1483 AACAGTTAGCAGAAGAAATTTAGTCAAGTCCAGTCCCAAGCTCAGAGTTCAGTGAACGT 1542
QY 461 AsnAspAspArgSerGluGluGlyLysTyrThrAlaValGlnArgAsnSerSerGluArg 480
DB 1543 AATGATGATAGGAGTGAGGAAGAAAAATACACAGCAGTTTCAGAGAAATTCAGTGAACGT 1602
QY 481 GluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGlyGlnArgAsnArg 500
DB 1603 GAGGGGCACAGCATAAACACATAGGAAAAATAAATATATATTCCTCTCGGCAAAAGAAAGA 1662
QY 501 GluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetGlyGlnProGlySer 520
DB 1663 GAATCATATCTCGGGAAGTGGGAGACAGAAATTCACCGCGTATGGCGCACCGCTGGATCG 1722
QY 521 GlySerMetProSerArgSerThrSerHisThrSerAspPheAsnProAsnSerGlySer 540
DB 1723 GGCTCCATGCCATCAAGATCCACTTCTCACACTTCAGATTTCAACCCGCAATTTCTGGTTCA 1782
QY 541 AspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerProSerSer 560

Db 1783 GACCAAGAGTAGTAAATGGAGGTTCCTCCGTGGCCATCGCCTTGCCCATCTCCTTCCTCT 1842
Qy 561 ArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProArgAlaAlaThrPro 580
Db 1843 CGCCCACTCTCTCGCTACCAAGTCAGTCCCAACTCTCTCCACCTCGGCGCAGCCACCCCT 1902
Qy 581 ThrArgProProSerArgProProSerArgProSerArgProProSerHisProSerAla 600
Db 1903 ACACGGCGCCCTCCAGGCGCCCTCGGCGCCATCCAGACCCCGCTCTCACCCCTCTGCT 1962
Qy 601 HisGlySerProAlaProValSerThrMetProLysArgMetSerSerGluGlyProPro 620
Db 1963 CATGGTCTTCCAGCTCTCTCTACTATGCTTAACCGCATGCTCTCAGAAAGGCGCTCA 2022
Qy 621 ArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGly 640
Db 2023 AGGATGTCCTCCCAAGGCGCCAGCGACATCCTCGAAATCACAGAGTTTCTGCTGGGAGGGT 2082
Qy 641 SerIleSerSerGlyLeuGluPheValSerHisAsnProProSerGluAlaAlaThrPro 660
Db 2083 TCCATATCCAGTGGCTAGAAATTTGTATCCCAACCAACCCAGTGAAGCAGCTACTCCT 2142
Qy 661 ProValAlaArgThrSerProSerGlyGlyThrTyrSerSerValValSerGlyValPro 680
Db 2143 CCAGTAGCAAGGACGAGTCCCTCGGCGGGAACGTGTCTATCATAGTGTCTAGTGGGTTTCA 2202
Qy 681 ArgLeuSerProLysThrHisArgProArgProArgSerProArgGlnAsnSerIleGlyAsnThr 700
Db 2203 AGATTATCCCTTAAACTCATAGACCCAGTCTCCAGACAGAACAGTATTGGNAATACC 2262
Qy 701 ProSerGlyProValLeuAlaSerProGlnAlaGlyIleProThrGluAlaValAla 720
Db 2263 CCCAGTGGGCCAGTCTTCTGCTTCTCCCAAGCTGGTATTATTCCAACTGAAGCTGTGGC 2322
Qy 721 MetProIleProAlaAlaSerProThrProAlaSerProAlaSerProAlaSerArgAlaValThr 740
Db 2323 ATGCGCTATTCCAGCTGCATCTCTAGCGCTGTAGTCTGATCGAATCGAACAGAGCTGTATCC 2382
Qy 741 ProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGly 760
Db 2383 CTTCTAGTAGGCTTAAAGATTCAGGCTTCAAGATCAGAGGCGAGAACTCTCTCGCAGG 2442
Qy 761 AsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLys 780
Db 2443 AATAAGAAATATTAACCCCAATGAACATCACCTAGCTTCTCAAAAGCTGAAACAAA 2502
Qy 781 GlyIleSerProValValSerGluHisArgLysGlnIleAspAspleuLysLysPheLys 800
Db 2503 GGTATATCACCAAGTCTTTCTGAACATGAACACAGATTGATTTAAAGAAATTTAAG 2562
Qy 801 AsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsnLys 820
Db 2563 ATGATTTTAGTTACAGCAAGTCTTCTCTGATCTATGATCAACTATAACAAA 2622
Qy 821 AsnArgGluGlyLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
Db 2623 AATAGAGGGGAGAAATCAAGAGATTTGATCAAAAGACAAAATTCAAACCAAGTGTCTAAG 2682
Qy 841 AspSerPheIleGluAsnSerSerAsnCysThrSerGlySerSerLysProAsnSer 860
Db 2683 GATTTCTTTTCAATGAAATAGCAGCAACTGTACCAGTGGCAGCAGCGCGAATAGC 2742
Qy 861 ProSerIleSerProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThr 880
Db 2743 CCAGCATTTCCCTTCAATCTATTAGTAACGAGACACAAGAGGGGAGCTGAGGTCACT 2802
Qy 881 SerGlnGlyValGlnThrSerSerProAlaCysLysGlnGluLysAspLysGluGlu 900
Db 2803 TCCCAAGGGGTTCCAGCTTCCAGCCCGCAGCATGTAAACAGAGAAAGACGATGAAGGAGAG 2862
Qy 901 LysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 2863 AAGAAAGACGAGCTGAGCAAGTTAGGAAATCAACATTGAATCCCAATGCAAGGAGGTTC 2922

Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 2923 AACCCAGTTCCTTCTCTCAGCAAAAGCTTCTACTACCCCAACTTCACTCGGCTCAA 2982
Qy 941 AlaGlnProSerProSerMetValGlyHisGlnGlnProThrProValTyrThrGlnPro 960
Db 2983 GCACAACTTACCCATCTATGGTGGTTCATCAACAGCCAACTCCAGTTTATCTCAGCT 3042
Qy 961 ValCysPheAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeu 980
Db 3043 GTTTGTTTTCACCAAAATATGATGTATCCAGTCCAGTGAGCCAGGCGTGAACCTTTA 3102
Qy 981 TyrProIleProMetThrProMetProValAsnGlnAlaLysThrTyrArgAlaValPro 1000
Db 3103 TACCAATATCCTATGACGCCCATGCCAGTGAATCAAGCAAGACATATAGAGCAGTACCA 3162
Qy 1001 AsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAlaMetMetHisProAlaSer 1020
Db 3163 AATATGCCCAACAGCGCAAGACCAAGCATCATCAGAGTGCATGATGCACCCAGCTCA 3222
Qy 1021 AlaAlaGlyProProIleAlaAlaThrProProAlaTyrSerThrGlnTyrValAlaTyr 1040
Db 3223 GCAGCGGCGCCACCGATTGCGAGCCACCCACAGCTTACTCCAGCAATATGTTGCCTAC 3282
Qy 1041 SerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGln 1060
Db 3283 AGTCTCAGCAGTTCGCCAAATCAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAG 3342
Qy 1061 HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThr 1080
Db 3343 CATCTCATGTCTATAGTCTCTGTAATACAGGTAATGCTAGAAATGATGCCACCAACA 3402
Qy 1081 HisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyValHisGluGlnThr 1100
Db 3403 CACGCCCGAGCTGGTTTAGTATCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGCG 3462
Qy 1101 HisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyr 1120
Db 3463 CATGGATGTATGCTATGCTCCCAATTTACATACAAAGGAGACAGCCCTCTCTTCTAC 3522
Qy 1121 PheAlaIleSerThrGlySerLeuAlaGlnTyrAlaHisProAsnAlaThrLeuHis 1140
Db 3523 TTTGCGATTTCCAGCGGCTCTCTGCTCAGCAGTATGCGACCTTAACGCTACCTCTGCAC 3582
Qy 1141 ProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHis 1160
Db 3583 CCACATATCTCCACACCTTCAGCTTACCTACCCCTGAGACAGCAGCAAGCCCAACAT 3642
Qy 1161 GlyGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAlaGlnAla 1180
Db 3643 GGTGAAGTATCTCTGCCACCCAGTCTGTTTACGACCATCAGCAGCCAGCCCGCCAGGCT 3702
Qy 1181 LeuHisLeuAlaSerProGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThr 1200
Db 3703 CTCCATCTGGCAGTCCACAGCAGCAGTCCATTTTACACGCGGGCTTGGCCCACT 3762
Qy 1201 ProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAlaAla 1220
Db 3763 CCACCTTCCATGACACCTTGCCTCCAAACAGCAGTCCGACACAGAAATAGTTTCCAGCAGCA 3822
Qy 1221 GlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnProPro 1240
Db 3823 CAACAGACTGTCTTTACGATCTCTTCTCAGCTTACGCGCGGTATACCAACCCACCC 3882
Qy 1241 HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSerHisProThr 1260
Db 3883 CACATGGGCCAGTACTCAGGCTCATGTACAGTCAAGAAATGTTCTTCTCATCAACT 3942
Qy 1261 AlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeu 1280
Db 3943 GCCCATGCGCCCAATGATGCTAATGACGACAGCAGCCCGCGGTCCCCAGGCCGCCCTC 4002

QY 1281 AlaGlnSerAlaLeuGlnProIleProValSerThrThrAlaHisPheProTyrMetThr 1300
 |||||
 Db 4003 GCTCAAGTGCCTACAGCCCATTCAGTCTCGACAAGCGCATTTCCCTTATATGACG 4062
 |||||
 QY 1301 HisProSerValGlnAlaHisGlnGlnGlnLeu 1312
 |||||
 Db 4063 CACCCTTCAGTACAAGCCACCACCAACAGCAGTTG 4098
 |||||

RESULT 5
 US-10-141-541-3
 ; Sequence 3, Application US/10141541
 ; Publication No. US20030167495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pulst, Stefan M.
 ; TITLE OF INVENTION: SCA2 Knockout Animal and Methods of Use
 ; FILE REFERENCE: P-CE 5244
 ; CURRENT APPLICATION NUMBER: US/10/141,541
 ; CURRENT FILING DATE: 2002-05-07
 ; PRIOR APPLICATION NUMBER: US 60/289,231
 ; PRIOR FILING DATE: 2001-05-07
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 4225
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (27)...(3884)
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (222)...(224)
 ; OTHER INFORMATION: cgs=Pro

US-10-141-541-3

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0	6025.50	4225	1159
Percent Similarity:	91.14%	Conservative:	45
Best local Similarity:	87.74%	Mismatches:	72
Query Match:	86.56%	Indels:	45
DB:	6	Gaps:	14

US-10-802-228-2 (1-1312) x US-10-141-541-3 (1-4225)

QY 1 MetArgSerAlaAlaAlaProArgSerProAlaValAla---ThrGluSerArgArg 19
 |||||
 Db 27 ATGCGTTGTCACCGCGCGGTTTCAGCGGCCCGCGCGGGGACCCCGAGCGCGCGCGC 86
 |||||
 QY 20 PheAlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgSerGly 39
 |||||
 Db 87 -----CGGCGGGCTGGCGCGCGCGGCTCGCTCCCGCGAGCGCGCGCGCGCGG 140
 |||||
 QY 40 ArgGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGly 59
 |||||
 Db 141 CGGGCGCGCGGTGGCG-----TATCCCTCCGCGCGCCCTCCCGCGCGCGC 188
 |||||
 QY 60 ProGlyProProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySer 79
 |||||
 Db 189 CCGCGCGCCCTCCCGCGCGCGGCTCGCCACCSTGCGGCTCAGACTGTGTTGGTAGC 248
 |||||
 QY 80 AsnGlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGlyGly 99
 |||||
 Db 249 AACGGCCAC-----GGGCGGTCCCGGCGCGGCTCCCGCGCGCTCGCTCGGTGTCNGCGG 302
 |||||
 QY 100 ProProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAla 119
 |||||
 Db 303 CTCCCGCGCCCTTCGTGTTGCTGCTCGCTCG-----GCCCGCGCGCGC 350
 |||||
 QY 120 AlaProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyVal--- 138
 |||||
 Db 351 AGCGCGCGCGCGCTGCGCGCGCGCGGCTCGCGCGCTCCCGCGCGCTCCGGCGTCTCC 410
 |||||

QY 139 SerLeuAlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGlyPro 158
 |||||
 Db 411 TCTCTGGCGCGCCCGGACCCCGGTCTCCCGCGCGCGTTCGAGCGCGTGTATGGGCGC 470
 |||||
 QY 159 LeuThrMetSerLeuLeuProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 178
 |||||
 Db 471 CTCACCATGTCGTGAAGCCGACGCCGAG----- 500
 |||||
 QY 179 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 198
 |||||
 Db 501 -----CCGCGCGCGCGCCGACTGGCGCGCAAGCCC 530
 |||||
 QY 199 GlyGlySerGlyLeuLeuAlaSerProAlaAlaProSerProSerSerVal 218
 |||||
 Db 531 GCGGCG---GGCCTGCTCTCGTCCCGCGCGCGCGCGC---TCGCGCGCGGTGACC 584
 |||||
 QY 219 SerSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSer-----Gly 235
 |||||
 Db 585 TCGGCTTCCGTTGCGCGCGCGCGCGCGCGCGCGTCTCTCGCGCGCGCGCGCGC 644
 |||||
 QY 236 GlyGlyArgProGlyLeuGlyArgGlnSerAsnGlyLeuProGlnSerThr 255
 |||||
 Db 645 GCGGCGGCTCCCGCGCTGGGCGAGGTCGGAACTAGCAAGAGACTGCCTCAGCCTACG 704
 |||||
 QY 256 IleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSerValVal 275
 |||||
 Db 705 ATTTCTTTTGATGGAATCTATGCAACGTGAGGATGGTTTCATATCTACGTCAGTTGTT 764
 |||||
 QY 276 GlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThr 295
 |||||
 Db 765 GGATCGAAATGTGAAGTACAAAGTGAAACCGGAGCATATATGAAGGAGTTTAAAAACA 824
 |||||
 QY 296 TyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSer 315
 |||||
 Db 825 TACAGTCTTAAGTGTGACTTGGTACTTGATGCTGCACATGAGAAAAGTACAGATCCAGT 884
 |||||
 QY 316 SerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAspPheVal 335
 |||||
 Db 885 TCGGGCGCAAAACGTGAAGAAATTAATGGAGAGTGTGTTTCAAATGCTCAGACTTCGTT 944
 |||||
 QY 336 ValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSer 355
 |||||
 Db 945 GTGGTACAGTTTAAAGATACAGACTCCAGTTTATGCACGAGAGATGCTTTTACTGACTCT 1004
 |||||
 QY 356 AlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAla 375
 |||||
 Db 1005 GCTCTCAGCGCAAGGTGAATGGTGAGCACAAAGGAGGACCTGGAGCCCTGGGATGCA 1064
 |||||
 QY 376 GlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTrp 395
 |||||
 Db 1065 GGGGAGGCTCAGCGCGAGCGAGCTGGAG---CTGGAGAAATGATGTGTCTTAATGGATGG 1121
 |||||
 QY 396 AspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAsp 415
 |||||
 Db 1122 GACCCCAATGACATGTTTCGATATAATGAAGAGAAATATGTTGGTGTCCACATATGAT 1181
 |||||
 QY 416 SerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLys 435
 |||||
 Db 1182 AGCAGTTTATCTTCATATACGTTCTTTAGAAAGGGACAACTCAGAAGAAATTTCTTAA 1241
 |||||
 QY 436 ArgGluAlaArgAlaAsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAla 455
 |||||
 Db 1242 CGGAGGCGAGGGCAACCACTAGTAGCAAGAAATTGAATCCAGTGTCTCAGTACAAAGCT 1301
 |||||
 QY 456 ArgValAlaLeuGluAsnAspAspArgSerGluGluLysTyrThrAlaValGlnArg 475
 |||||
 Db 1302 CGTGTGCGCCCTTGAGATGATGACCGGAGTGAGGAAGAAAATACACAGCAGTCCAGAGA 1361
 |||||
 QY 476 AsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProPro 495
 |||||
 Db 1362 AACTGCAGTACCGGGAGGGGCGATGGCCCAACACTAGGGACAATAAATATATTCTCTCT 1421
 |||||
 QY 496 GlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMet 515
 |||||

1422 GGACAAAGAAACAGAGAGTCTATCTCTGGGAAAGTGGAGACAGAGCTCAACCGGATG 1481
516 GlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAspPheAsn 535
1482 GGCCAGCCTGGGCCAGGCTCCATGCGGTCAAGAGCTGCTTCTCACACTTCAGATTCAAC 1541
536 ProAsnSerGlySerAspGlnArgValValAsnGlyValProTrpProSerProCys 555
1542 CCGAAGCTGGCTCAGACCAAGAGTAGTTAATGAGGTGTTCCCTGGCCATCGCCTTGC 1601
556 ProSerProSerSerArgProProSerArgProSerArgProSerArgProPro 575
1602 CCATCTCATCTCTCGCCACCTTCTCGCTACCAAGTCAGGTCCCAACTCTCTTCAACT 1661
576 ArgAlaAlaThrProThrArgProProSerArgProProSerArgProSerArgProPro 595
1662 CGGCGAGCCACCCATACAGCGCGCCCTCCAGGCCCTCCAGGCCATCCAGACCCCG 1721
596 SerHisProSerAlaHisGlySerProAlaProValSerThrMetProLysArgMetSer 615
1722 TCTCACCCCTCTGCTCATGTTCTCCAGCTCCTGTCTCTACTATGCTTAACGCGATGCT 1781
616 SerGluGlyProProArgMetSerProLysAlaGlnArgHisProArgAsnHisArgVal 635
1782 TCAGAAAGGACCCCAAGAGTGTCTCCAAAGGCACAGCGCCACCTCGGAATCACAGAGTC 1841
636 SerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHisAsnProProSer 655
1842 TCTGCTGGGAGAGGCTCCATGCTAGTGCCCTAGAAATTTGTATCCCAATATCCCCCAAT 1901
656 GluAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThrTrpSerSerVal 675
1902 GAAGCAGCTCTCTCCAGTGGCAAGACCAAGTCTCTGCAGGGGAGGAGTGTCTCAGTG 1961
676 ValSerGlyValProArgLeuSerProLysThrHisArgProArgSerProArgGlnAsn 695
1962 GTCAGTGGGGTTCCAAGGTATCTCCCAAACTCACAGACCCCAAGGTCTCCAGGCAGAGC 2021
696 SerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAlaGlyIleIlePro 715
2022 AGCATTTGGAAACTCTCCAGCGGCGCTGTCTGCTTCTCCCAAGCTGCAATCATCCCT 2081
716 ThrGluAlaValAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSer 735
2082 GCAGAGCCGTTTCCATGCTGTTCCCGCGCATCTCCGACTCTCGCAGCCCTGCATCC 2141
736 AsnArgAlaValThrProSerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGln 755
2142 AACAGAGCACTGACCCCATCTATTGAGGCAAAAGATTCCAGGCTTCAAGATCAGAGGCAG 2201
756 AsnSerProAlaGlyAsnLysGluAsnIleLysProAsnGluThrSerProSerPheSer 775
2202 AACTCTCTCCAGGGAGTAAGAAAAATGTTAAAGCAAGTGAACATCACCCTAGCTTTTCA 2261
776 LysAlaGluAsnLysGlyIleSerProValValSerGluHisArgLysGlnIleAspAsp 795
2262 AAAGCTGACAAACAAAGTAGTGTCCACAGTGTGTTCTGAAACAGAAAAACAGATTGATGAC 2321
796 LeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAsp 815
2322 TTAAGAAGTTTAAGATGATTTTAGTTTACAGCAAGCTCTACATCTGAATCTATGGAT 2381
816 GlnLeuLeuAsnLysAsnArgGluGlyLysSerArgAspLeuIleLysAspLysIle 835
2382 CAACTACTAAGCAAAAAATAGAGAGGAGAAAGTCAACGAGATTGATTAAAGATAAAACG 2441
836 GluProSerAlaLysAspSerPheIleGlu-----AsnSerSerSerAsnCysThrSer 853
2442 GAACCAAGTCTAAGGATAGTTTTCATTGACAGCAGCAGCAGCAGCACTGTACCACT 2501
854 GlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnThrGluHis 873

2502 GGCAGCAGCAAGACCAACAGCCCTAGCATCTCCCTTCCATGCTTAGTAATGACAGCAC 2561
874 LysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProAlaCysLysGln 893
2562 AAGAGGGGCGCTGAGTGCATATCCAAAGGGTGCAGACTTCCAGCCAGCCCTGCAACAA 2621
894 GluLysAspAspLysGluGluLysLysAspAlaAlaGluGlnValArgLysSerThrLeu 913
2622 GAGAAGGATGACAGAGAGAGAAAGAACACAAACAGAGCAGGTTTGAAGAAATCGACATTG 2681
914 AsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProLysProSerThrThr 933
2682 AATCCCAATGCAAGAGGAGTTCAACCTCTGTTCTCTCAGCCAAAGCCCTTCTACTACC 2741
934 ProThrSerProArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGlnPro 953
2742 CCAACGTCACCTCGGCCCTCAAGCACCAACAGCCCATCTATGTTGGTGCATCAGCAGCCA 2801
954 ThrProValThrThrGlnProValCysPheAlaProAsnMetMetTyProValProVal 973
2802 GCTCCAGTGTACACTCAGCCTGTGTGCTTCGCACCAATATATGATATCCCTGCCAGTG 2861
974 SerProGlyValGlnProLeuTyProIleProMetThrProMetProValAsnGlnAla 993
2862 AGCCCGGGGCTACAACTTTATACCAATACCTATGACGCCCATGCTGTGAACCAAGCC 2921
994 LysThrTyArgAla-----ValProAsnMetProGlnGlnArgGlnAspGlnHisHis 1011
2922 AAGACATATAGCAGAGGTAAAGTACCAATATGCCCCAACAGCAGCAACCAACATCAT 2981
1012 GlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThrProPro 1031
2982 CAAAGCACCATGATGACCCAGCCTCCCGGAGGGCCACCCATCGTAGCCACCCTCCGCC 3041
1032 AlaTySerThrGlnTyValAlaTySerProGlnGlnPheProAsnGlnProLeuVal 1051
3042 GCTTACTCACTCAGTACGTTGCTTACAGCCCTCAGCAGTTCCTCCCAATCAGCCTTTGGTC 3101
1052 GlnHisValProHisTyGlnSerGlnHisProHisValTySerProValIleGlnGly 1071
3102 CAGCATGTGCCCATTTATCAGTCTCAGCATCTCTCATGTGTACAGTCTCTCATACAAGGT 3161
1072 AsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerAla 1091
3162 AATGCGAGGATGATGCGCACCAACAGCAGCATGCTCAGCCTGGTTAGTGTCTTCTCAGCT 3221
1092 ThrGlnTyGlyValHisGluGlnThrHisAlaMetTyAlaCysProLysLeuProTy 1111
3222 GCTCAGTTCGGGGCTCAGCAGCAGCAGCAGCAGCAGTGTATGATGTCTCCCAATTTACCATAC 3281
1112 AsnLysGluThrSerProSerPheTyPheAlaIleSerThrGlySerLeuAlaGlnGln 1131
3282 AACAGGAGACAAAGCCCTTCTTCTTACTTTGCCATTTCCACCGCTCCCTCGCTCAGCAG 3341
1132 TyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSerAlaThr 1151
3342 TATGCACATCTTAATCCGCCCTGTCATCCACATACCTCCCATCTCCTCAGCCTTCGGCCAT 3401
1152 ProThrGlyGlnGlnGlnSerGlnHisGlyGlySerHisProAlaProSerProValGln 1171
3402 CCACCGGACAGCAGCAAGCCAGCATGTTGGAAGTCACTCCCTGCACCCAGTCTCTGTTTCAG 3461
1172 HisHisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerProGlnGlnGlnSerAla 1191
3462 CACCATCAGCACCGGCTGCCAGGCTCTTCATCTGCCAGTCCACAGCAGCAGTCCGCC 3521
1192 IleTyHisAlaGlyLeuAlaProThrProProSerMetThrProAlaSerAsnThrGln 1211
3522 ATTTATCATCGGGGCTGGCACCAACCAACCTTCCATGACACCTGCTCTTAATACACAG 3581
1212 SerProGlnAsnSerPheProAlaAlaGlnThrValPheThrIleHisProSerHis 1231
3582 TCTCCACAGCAGCAGTTCCTCCAGCAGCAGCAACAGACAGTCTTCCACCATCCACCTTCTCAT 3641


```
QY 1232 ValGlnProAlaTyrThrAsnProProHisMetAlaHisValProGlnAlaHisValGln 1251
Db |||||
Db 3642 GTTCAGCGGCATACACACCCACCCACATGGCCACGTACCTACCTCAGGCTCATGTACAG 3701
QY 1252 SerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeuMetThrThrGln 1271
Db |||||
Db 3702 TCAGAAATGGTTCTCTCATCACTGCCATGCCCAATGGCCAATGATGCTAATGACGACACAG 3761
QY 1272 ProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnProPheProValSer 1291
Db |||||
Db 3762 CCACCC--GGTCCCAAGGCCGCTCGCTCAAGTGCACTACACCCATTCACGTTTCG 3818
QY 1292 ThrThrAlaHisPheProTyrMetThrHisProSerValGlnAlaHisGlnGlnGln 1311
Db |||||
Db 3819 ACAACAGCGCATTTCCCTTATATGACGACCCCTTCAGTACAGGCCACCAACACAGCAG 3878
QY 1312 Leu 1312
Db |||
Db 3879 TTG 3881

RESULT 6
; Sequence 15, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 15
; LENGTH: 2006
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (862)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (1006)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-298-15

Alignment Scores:
Pred. No.: 1,42e-138 Length: 2006
Score: 2463.00 Matches: 460
Percent Similarity: 98.72% Conservative: 1
Best Local Similarity: 98.50% Mismatches: 2
Query Match: 35.38% Indels: 4
DB: 3 Gaps: 1

US-10-802-228-2 (1-1312) x US-09-925-298-15 (1-2006)
QY 850 AenCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSer 869
Db : : :
Db 6 AGCTGTACAGTGGCAGCAGCAAGCGGAATAGCCCCAGCATTTCCCTTCAATACTTAGT 65
QY 870 AenThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerPro 889
Db |||||
Db 66 AACACGGAGCACAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTTCAGACTTCCAGCCCA 125
QY 890 AlaCysLysGlnGlyAspAspLysGluGluLysLysAspAlaAlaGlnValArg 909
Db |||||
Db 126 GCATGTAAACAGAGAGACGATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 185
QY 910 LysSerThrLeuAsnProAsnAlaAlaLysGluPheAsnProArgSerPheSerGlnProLys 929
Db |||||
```

```
Db 186 AAATCAACATTGAATCCCAATGCAAGGAGTTC AACCCACGTTCTCTCTCAGCCAAAG 245
QY |||||
QY 930 ProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetValGly 949
Db |||||
Db 246 CTTTCTACTACCCCAACTTCACCTCGGCCTCAAGCACAACTAGCCCATCTATGGTGGGT 305
QY |||||
QY 950 HisGlnGlnProThrProValTyrThrGlnProValCysPheAlaProAsnMetMetTyr 969
Db |||||
Db 306 CATCRACAGCCCACTCCAGTTTATACCTGCTGTTGTTTGGACCAAAATATGATGTAT 365
QY |||||
QY 970 ProValProValSerProGlyValGlnProLeuTyrProLeuProMetThrProMetPro 989
Db |||||
Db 366 CCAGTCCAGTCAGCCAGCGGTGCAACTTTATACCCAATACCTATGACGCCCATGCCA 425
QY |||||
QY 990 ValAsnGlnAlaLysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGln 1007
Db |||||
Db 426 GTGAATCAAGCCCAAGACATATAGAGCAGGTAAAGTACCAAAATATGCCCAACAGCGGCA 485
QY |||||
QY 1008 AspGlnHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAla 1027
Db |||||
Db 486 GACCAGCATCATCAGAGTGCCATGATGCACCCAGCGTCAGCAGCGGCCCCACCGATTGCA 545
QY |||||
QY 1028 AlaThrProProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsn 1047
Db |||||
Db 546 GCMACCCCAACAGCTTACTCCACGCAATATGTCCTACAGTCTCCTCAGCAGTTCCCAAT 605
QY |||||
QY 1048 GlnProLeuValGlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerPro 1067
Db |||||
Db 606 CAGCCCTTGTTCAGCATGTGCCACATTTATCAGTCTCAGCATCTCATGTCTATAGTCT 665
QY |||||
QY 1068 ValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuVal 1087
Db |||||
Db 666 GTAATCAGGGTAATGCTAGTAATGATGCGACCAACACACAGCGGCCCGCTGGTTAGTA 725
QY |||||
QY 1088 SerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysPro 1107
Db |||||
Db 726 TCTTCTTCAGCAACTCAGTACGGGGCTCATGAGCAGACGCGATGCGATGATGATGTC 785
QY |||||
QY 1108 LysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySer 1127
Db |||||
Db 786 AAATTACCATACAAAGGAGACAAAGCCCTTCTTCTACTTTGGCATTTCCACGGGCTCC 845
QY |||||
QY 1128 LeuAlaGlnGlnTyr-AlaHisProAsnAlaThrLeuHisProHisThrProHisProGln 1147
Db |||||
Db 846 CTTCCTCAGCAGTATGNGCRACCTTACGCTACCTGCACCCACATACTCCACACCCCTCA 905
QY |||||
QY 1147 nProSerAlaThrProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaPr 1167
Db |||||
Db 906 GCCTTCAGCTACCCCACTGGACAGCAGCAAGCAACATGGTGGAAAGTCATCTCTGCACC 965
QY |||||
QY 1167 oSerProValGlnHisGlnHisGlnAlaAlaGlnAla-LeuHisLeuAlaSerProG 1187
Db |||||
Db 966 CAGTCTCTGTTTCAGCACCATCAGCAGCGCCAGCGCTTCTCATCTGCGCCAGTCCAC 1025
QY |||||
QY 1187 InGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProProSerMetThrProA 1207
Db |||||
Db 1026 AGCAGCAGTCAAGCCATTTACACCGGGGCTTGGCCCAACTCCACCCTCCATGACACCTG 1085
QY |||||
QY 1207 laSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrI 1227
Db |||||
Db 1086 CCTCCAAACAGCAGTCGCGCAGAAATAGTTTCCAGCAGCAGCAACAGAGACTGTYTTACGA 1145
QY |||||
QY 1227 leHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisValProG 1247
Db |||||
Db 1146 TCCATCTCTTCAGCTTTCAGCGCGGTATACCAACCCACCCACATGGCCCGCTTACCTC 1204
QY |||||
QY 1247 InAlaHisValGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetL 1267
Db |||||
Db 1206 AGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAACTGCCCATCGCGCAATGATGC 1265
QY |||||
QY 1267 euMetThrThrGlnProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnP 1287
Db |||||
```

Db 1266 TAATGACGACACAGCCACCGGGCTCCAGCGCGCCCTCGCTCAAAAGTGCACTACAGC 1325
 QY 1287 roileProValSerThrThrAlaHisPheProTyrMetThrHisProSerValGlnAlaH 1307
 Db 1326 CCATTCCAGTCTCGACCAACAGCGGCATTTCCCTATATGACGACCGCTTCAGTACAAGCCC 1385
 QY 1307 isHisGlnGlnGlnLeu 1312
 Db 1386 ACCACCAACAGCAGTTG 1402
 RESULT 7
 US-10-102-806-15
 ; Sequence 15, Application US/10102806
 ; Publication No. US20030054421A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA103PIC1
 ; CURRENT APPLICATION NUMBER: US/10/102,806
 ; CURRENT FILING DATE: 2002-03-22
 ; PRIOR APPLICATION NUMBER: 09/925,298
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05881
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 846
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 15
 ; LENGTH: 2006
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (862)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: misc_feature
 ; LOCATION: (1006)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-10-102-806-15
 Alignment Scores:
 Pred. No.: 1,42e-138 Length: 2006
 Score: 2463.00 Matches: 460
 Percent Similarity: 98.72% Conservative: 1
 Best Local Similarity: 98.50% Mismatches: 2
 Query Match: 35.38% Indels: 4
 DB: 5 Gaps: 1
 US-10-802-228-2 (1-1312) x US-10-102-806-15 (1-2006)
 QY 850 AsnCysThrSerGlySerLysProAsnSerProSerIleSerProSerIleLeuSer 869
 Db 6 AGCTGTACAGTGGCAGCAGCAGCGAATAGCCCGCAGCATTTCCCTTCATACTAGT 65
 QY 870 AsnThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerPro 889
 Db 66 AACACGGAGCACAAAGAGGGGACCTGAGGTCACTTCCCAAGGGGTTCCAGACTTCCAGCCCA 125
 QY 890 AlaCysGlyGlnLysAspAspLysGluLysLysAspAlaGluGlnValArg 909
 Db 126 GCATGTAAACAGAGAAAGACGATAGGAAGAGAGAAAGACGCGCTGAGCAAGTTAGG 185
 QY 910 LysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSerPheSerGlnProLys 929
 Db 186 AAATCAACATTGATCCCATGCAAGAGAGTTCAACCCAGTTCTCTCTCAGCCCAAG 245
 QY 930 ProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetValGly 949
 Db 246 CCTTCTACTACCCCAACTTCACCTCGCGCTCAAGCAACACCTAGCCCATCTATGGTGGGT 305
 QY 950 HisGlnGlnProThrProValThrGlnProValCysPheAlaProAsnMetMetTyr 969

Db 306 CATCAACAGCCAACTCCAGTTTATATACTCAGCGTGTGTTGTTTGCACCAAAATATGATGAT 365
 QY 970 ProValProValSerProGlyValGlnProLeuTyrProIleProMetThrProMetPro 989
 Db 366 CCAGTCCAGTGGAGCCCGAGCGGTGCAACCTTTATACCAATACCTTATGAGCCCATGCCA 425
 QY 990 ValAsnGlnAlaLysThrTyrArgAla-----ValProAsnMetProGlnGlnArgGln 1007
 Db 426 GTGAATCAAGCCAGACATATAGAGCAGGTAAGATACCAAAATATGCCCAACAGCGGCAA 485
 QY 1008 AspGlnHisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAla 1027
 Db 486 GACCAGCATCATCAGTGCATGATGACACCGCTCAGCAGCGGGCCACCGATTGCA 545
 QY 1028 AlaThrProProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPheProAsn 1047
 Db 546 GMAACCCACCGAGCTTACTCCACGCAATATGTTCCTACAGTCTCAGCAGTTCCCAAA 605
 QY 1048 GlnProLeuValGlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerPro 1067
 Db 606 CAGCCCTTGTTCAGCATGTGCCACATTATCAGTCTCAGCATCTCATGTCTATAGTCTC 665
 QY 1068 ValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuVal 1087
 Db 666 GTAAATACAGGGTAATGCTAGAAATGATGACCAACACACACCGCCAGCTCGTTTAGTA 725
 QY 1088 SerSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysPro 1107
 Db 726 TCTTCTCAGCAACTCAGTACGGGCTCATGAGCAGACGCGATGCGATGATCATGTGTC 785
 QY 1108 LysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySer 1127
 Db 786 AAATACCACATACCAAGAGAGACAGCCCTCTTTCTTACTTTGCCATTTCCACGGGCTCC 845
 QY 1128 LeuAlaGlnGlnTyr-AlaHisProAsnAlaThrLeuHisProHisThrProHisProG 1147
 Db 846 CTTGCTCAGCAGTATGNGCRACCTTAACGCTACCTCGACCCACACACTACTCCACACCTCA 905
 QY 1147 nProSerAlaThrProThrGlyGlnGlnSerGlnHisGlySerHisProAlaPro 1167
 Db 906 GCCTTCAGCTACCCCACTGGGACAGCAAGCAAGCAACATGTTGGAAAGTCATCTGCACC 965
 QY 1167 OSerProValGlnHisHisGlnAlaGlnAlaGlnAla-LeuHisLeuAlaSerProG 1187
 Db 966 CAGTCTGTTCAGCAGCAGCAGCAGCGCCCGCCAGGCTNCTCCATCTGGCCAGTCCAC 1025
 QY 1187 InGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProSerMetThrProA 1207
 Db 1026 AGCAGCAGTCAGCCATTTACCAACGCGGGCTTGGCCCAACTCCACCTCCATGACACCTG 1085
 QY 1207 laSerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrI 1227
 Db 1086 CCTCCAACAGCAGTCGCCACAGAAATAGTTTCCAGCAGCAGCAACAGACTGTGTTTTACGA 1145
 QY 1227 leHisProSerHisValGlnProAlaTyrThrAsnProProHisMetAlaHisValProG 1247
 Db 1146 TCCATCTCTTCTCAGCTTACGCGGGGTATACCAACCCACCCAGCTGGCCCGCAGTACCTC 1205
 QY 1247 InAlaHisValGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetL 1267
 Db 1206 AGGCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAACTGCCCATGCGCAATGATGC 1265
 QY 1267 euMetThrThrGlnProProGlyGlyProGlnAlaLeuAlaGlnSerAlaLeuGlnP 1287
 Db 1266 TAATGACGACACAGCCACCGCGGTCCCGCGGCGCTCGCTCAAGTGCACTACAGC 1325
 QY 1287 roileProValSerThrThrAlaHisPheProTyrMetThrHisProSerValGlnAlaH 1307
 Db 1326 CCATTCCAGTCTCGACCAACAGCGGCATTTCCCTATATGACGACCGCTTCAGTACAAGCCC 1385
 QY 1307 isHisGlnGlnGlnLeu 1312
 Db 1386 ACCACCAACAGCAGTTG 1402


```

RESULT 10
US-10-171-581-346
; Sequence 346, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mac Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171.581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 346
; LENGTH: 4674
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AF034373
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-346

```

Alignment Scores:

```
Pred. No.: 2,16e-99 Length: 4674
Score: 1816.50 Matches: 504
Percent Similarity: 48.22% Conservative: 174
Best Local Similarity: 35.85% Mismatches: 381
Query Match: 26.10% Indels: 350
DB: 5 Gaps: 52

US-10-802-228-2 (1-1312) x US-10-171-581-346 (1-4674)

QY 19 ArgPheAlaAlaAlaAArgTProGlyTyrArgSerLeuGlnArgProAlaArgSer 38
DB 6 CGCTCTCACTTCTAGATGGATGGCGCGCGGAGAGGCGCTCTCTATTCTTAGATG 65
QY 39 GlyArgGlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProPro 58
DB 66 GATGGCGCGCGGCGAGAGATGCTCTCACTTCTCCAGCGCGAGACCCCTCCCTT 125
QY 59 GlyProGlyProProProSerArgGlnSer-----SerProSerAlaSerAspCys 76
DB 126 CGCTCTCGCGCGGCTTCTCGCGCGCGGCTTCTCTCTCACCC-----170
QY 77 PheGlySerAsnGlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuGly 96
DB 171 -----CCGACACCGCG-----182
QY 97 LeuGlyGlyProProArgProPheValValLeuLeuLeuProLeuAlaSerProGlyAla 116
DB 183 -----GGCTCCCGCGCGCG-----CCGACGCGCGCGCGCGCTGC 218
QY 117 ProProAlaAlaProThraArgAlaSerProLeuGlyAlaArgAlaSerProProArgSer 136
DB 219 CGATCCCGCTCGCTTCCCGCGC-TCTCCAGCGGCGCGCCAGCCCGCGCTCTCTCC 277
QY 137 GlyValSerLeuAlaArgProAlaProGlyCysProArgProAlaCysGluProValTyr 156
DB 278 CTCCTCTCTCTA-ATTCCCTTCCGACGCT-----309
QY 157 GlyProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGln 176
DB 310 -----GCCATCATGTTGAAGCCTCAGCGCTTACACAGCCCTCCAGCGCGCGCAG 360
QY 177 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
DB 361 CGCCCGCCACCAACATGCGTGGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 196 GlySerProGlyGly-----SerGlyLeuLeu 204
DB 421 CGGGCTCCGGGGCGCTGGCCACCTCTGGCGCTCTCCGGGCGCTCCAGCGCGCGCT 480
QY 204 uAlaSerProAlaAlaAla-----ProSerProSerSerSerValse 219
DB 481 CCCCTCGCTGGGGCTGTGGCGCTGGCGGAGCGGCTCCGCGGGAGCGCGAGGCA 540
QY 219 rSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGlyArgPr 239
DB 541 TCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 239 o-GlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSerPheA 259
DB 601 TCGGAGCGCGCGGAGCAGACGAGGAAGAGGACCCCGACGTCACCTGTG---TTG 657
QY 259 spGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySerLysC 279
DB 658 AAGCGCTTACAACAATTCAGAAATGCTGATTTCTTACAGCTGTGTGGGCTCCACTT 717
QY 279 ysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProL 299
DB 718 GTGATGTAAGGTGAAAGTATGTTACCTTATGAGGGTATCTTCAAGACGCTAAGCTCAA 777
QY 299 ysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerGlyProL 319
DB 778 AGTTGAACCTAGCGGTGATGCTGTGCACCGGAAGACATCTGAGCCAGCGTGGCGCTC 837
```

```
QY 319 ysArgGluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnP 339
DB 838 CTCGGGAGGACATTTGTGACACCATGGTGTAAAGCAAGTCAATGCTGTTGTTCACT 897
QY 339 heLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIle--- 357
DB 898 TCCGAAATGTTGACTTCAACTATCTACTAAAGACAAGTTCACCGATTCAACCATTTGCCA 957
QY 358 --SerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAlaGlyG 377
DB 958 TGAACCTCGAAAGTGAATGGGGAACACAAGAGAGAGGTCTTTCAGCGCTGGAGGGGGTG 1017
QY 377 luLeuThrAlaAsnGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTyrAspP 397
DB 1018 AC-----AGCAACAGCGACGACTATGACCTCGAGTCTGACATGTCCAATGGATGGGACC 1071
QY 397 roAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAspSerS 417
DB 1072 CCAATGAATGTTCAAGTTCATGAGGAGAACTACGGTGTGAAGACTACCTATGATAGCA 1131
QY 417 erLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLysArg 437
DB 1132 GTCCTTCTTCTATACGGTGCCTTAGAAAGAGCACTCAGAACAGTTCCTGACGCGAG 1191
QY 437 luAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAlaArgV 457
DB 1192 AGCTGCGTGGCGCGCGCTGCTCGAGAGATGAATCAAGCCCCCAGTACCGCTACGGA 1251
QY 457 alAlaLeuGluAsnAspAsp---ArgSerGluGluGluLysTyrThrAlaValGlnArgA 476
DB 1252 TCGCATGAGAGAACGACGATGGCGCCTGAAAGAGGAGAACACAGTCAGTCCAGCGCGC 1311
QY 476 snSerSerGluArgGlyGlyHisSerIleAsnThrArgGluAsnLysTyrIleProArg 496
DB 1312 AGGCTCAGCGCGGAGAGCGCCAGCTTGGCATCCAGGAGGGAGATATATCTCTCTGC 1371
QY 496 lyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMetG 516
DB 1372 CTCACGAGTCCGGGAAGGT---CCCCGGGAGGAGTTCGATCGACGAGCTCTCGGGCGC 1428
QY 516 lyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAspPheAnp 536
DB 1429 GTCGCGCTGGCTTGTAGCTCTTTGCGCCTCTGCGCCCTCACCATCTGGACACAGCAGCC 1488
QY 536 roAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTrpProSerProCysP 556
DB 1489 CTGGCCCGAGTTCAGCGCGCTGGTATCAATGGA-----1523
QY 556 roSerProSerSerArgProProSerArgTyrGlnSerGlyProAsnSerLeuProProA 576
DB 1523 -----1523
QY 576 rgAlaAlaThrProThrArgProProSerArgProProSerArgProSerArgProProS 596
DB 1523 -----1523
QY 596 erHisProSerAlaHisGlySerProAlaProValSerThrMetProLysArgMetSerS 616
DB 1523 -----1523
QY 616 erGluGlyProProArgMetSerProLysAlaGlnArgHisProArgAsnHisArgValS 636
DB 1524 -----GGCCCTTCCCGCATGTCCCAAGAGGACAGCGCGCTCTGAGA-----1565
QY 636 erAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHisAsnProProSerG 656
DB 1566 -----GGTGCCACAGACTCTGTCTTCG-----CCCACTA 1593
QY 656 luAlaAlaThrProProValAlaAlaArgThrSerProSerGlyThrTrpSerValV 676
DB 1594 ATAGGCTTCTGGAGAACTTCTGTCTCCACCTCTCTCTGAGCTGGG-----1637
```



```
Db      3528 ACATCGGACACCC 3541
RESULT 11
US-09-823-245A-600/c
; Sequence 600, Application US/09823245A
; Publication No. US20020039760A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fachtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6401
; CURRENT APPLICATION NUMBER: US/09/823,245A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/194,941
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 600
; LENGTH: 1707
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-823-245A-600

Alignment Scores:
Pred. No.:      2,328-53      Length:      1707
Score:          1042.50      Matches:    264
Percent Similarity: 52.05%      Conservative: 105
Best Local Similarity: 37.24%      Mismatches: 168
Query Match:      14.98%      Indels:     172
DB:                3          Gaps:          24

US-10-802-228-2 (1-1312) x US-09-823-245A-600 (1-1707)
QY      198  ProGlyGlySerGlyLeuAlaSerProAlaAlaProSerProSerSerSerSer 217
Db      1698  CTGGGGCTGTGGCGCTGGCGCGGAGCGGGCT-----CCGCGGGGAGCCGA 1651
QY      218  ValSerSerSerAlaThrAlaProSerValValAlaAlaThrSerGlyGlyGly 237
Db      1650  AGGCATCTTGGCGCGCAGCGCGCGCGCAGCAACACAGGAGAGCGCGGGGCGAGC 1591
QY      238  ArgPro-GlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSerThrIleSe 257
Db      1590  CGCCATCGGCAGCGCGGAGGAGCAGAGCAGAGAAAGGAGCCGCCACAGTCACCTGTG- 1533
QY      257  rPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSerValValGlySe 277
Db      1532  -TTTGAAGCGGTCTACAACAATTCAGAAATGCTGCAATTCCTTACAGCTGTGTGGGCTC 1474
QY      277  rLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSe 297
Db      1473  CACTTGTGATGTAAGGTGAAAAATGGTACCCTTATGAGGGTATCTTCAAGACGCTAAG 1414
QY      297  rProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGl 317
Db      1413  CTCAAAAGTTGAACCTAGCGGTGATGCTGTGCACCGGAAAGCATCTGAGCCAGCAGGTGG 1354
QY      317  yProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValva 337
Db      1353  CCCTCGTCGGGAGGACATGTGGACACCATGGTGTGTTAAGCAAGTGTATGTCCTTGT 1294
QY      337  lGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIle 357
Db      1293  TCACCTTCGGAATGTTGACTTCAACTATGCTACTTAAGACAGATTACCGATTACGCCAT 1234
QY      357  e-----SerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluProTrpAspAl 375
Db      1233  TGCCATGAACCTCGAAAGTGAATGGGAAACACAAAGAGAGAGGTGCTTCAGCGCTGGAGGG 1174
QY      375  aGlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAsnGlyTr 395
Db      1173  GGGTGAC-----AGCAACAGCGAGCACTATGACCTCGAGTCTGACATGTCCAATGGATG 1120
QY      395  pAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerThrTyrAs 415
Db      1119  GGACCCCAATGAAATGTTCAAGTTCAATGAGGAGAACTACGGTGTGAAGACACTACCTATGA 1060
QY      415  pSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPheLeuLy 435
Db      1059  TAGCAGTCTTTCTTATACGGTGCCTTGAAGAAGACCAACTCAGAAGAGATTTCTGTCA 1000
QY      435  sArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTyrLysAl 455
Db      999  GCGAGAGCTGGTGGCGGCCAGTTGGCTCGAGAGATTGAATCAAGCCCCCAGTACCGCT 940
QY      455  aArgValAlaLeuGluAsnAspAsp---ArgSerGluGluGluLysTyrThrAlaValGl 474
Db      939  ACGGATCGCCATGGAGAACGACGATGGCGGCACCTGAACAGAGAGAACACAGTGCAGTCCA 880
QY      474  nArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIlePr 494
Db      879  GCGGCAGGGCTCAGGGCGGAGAGCCCGAGCTGGCATCCAGGGAGGGGAAATATATCCC 820
QY      494  oProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProAr 514
Db      819  TCTGCTCAACGAGTCCGGGAAGGT---CCCCGGGGAGGAGTTTCATCAGCAGCTCTCG 763
QY      514  gMetGlyGlnProGlySerGlySerMetProSerArgSerThrSerHisThrSerAspPh 534
Db      762  GGGCGGTGGCGCTGGCCTTAGCTCTTTGCCACCTCGTGGCCCTCACCATCTGGACACAG 703
QY      534  eAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValProTrpProSerPr 554
Db      702  CAGCCCTGGCCCGAGGTTCTGAGGCCCGTGTGTATCAATGGA----- 663
QY      554  oCysProSerProSerSerArgProProSerArgProGlnSerGlyProAsnSerLeuPr 574
Db      663  ----- 663
QY      574  oProArgAlaAlaThrProThrArgProProSerArgProProSerArgProSerArgPr 594
Db      663  ----- 663
QY      594  oProSerHisProSerAlaHisGlySerProAlaProValSerThrMetProLysArgMe 614
Db      663  ----- 663
QY      614  tSerSerGluGlyProProArgMetSerProLysAlaGlnArgHisProArgAsnHisAr 634
Db      662  -----GGCCCTTCCCGCATGTCCCAAAGACACAGCGGCTCTGAGA----- 621
QY      634  gValSerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHisAsnProPr 654
Db      620  -----GGTGCCAGACTCTGTCTTCG-----CCACGTAATAGGCCTTC 583
QY      654  oSerGluAlaAlaThrProProValAlaAlaArgThrSerProSerGlyGlyThrTrpSerSe 674
Db      582  TGGAGAAACTTCTGTTCACCT-----CCTCTCGAGTGGGC----- 546
QY      674  rValValSerGlyValProArgLeuSerProLysThrHisArgProArgSerProArgGl 694
Db      545  -----CGGATGTATCCC-----CGCGCTTCTTCCCAAG-- 519
QY      694  nAsnSerIleGlyAsnThrProSerGlyProValLeuAlaSer-----ProGl 710
Db      518  -----TCTGTGGCCCTGCCCAATCTCAGCTTCTCTCCAGAGCCTCCCAT 472
QY      710  nAlaGlyIleIleProThrGluAlaValAlaMetProIleProAlaAla----- 726
Db      710  nAlaGlyIleIleProThrGluAlaValAlaMetProIleProAlaAla----- 726
```

Db 471 CGGCTGGGAGTGGCCAACTCTTACGCTCCATCCCTGTGACCTCATCAGTCTCAGATCC 412
QY 727 -----SerProThrProAlaSerProAlaSerProAlaSerProAlaValThrProSerSe 743
Db 411 TGGAGTGGGCTCCATTTCTCCAGTCTCTCCAAAGATCTCCCTGGCCCCCACA----- 360
QY 743 rGlualAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGl 763
Db 359 -----GATGTAAAGAACTCTTACCAAGAAACCTGGGAG 325
QY 763 uAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSe 783
Db 324 AACTCTGGAGCCCGAGAGCTG-----GCTCGATAGCTGGGAAGT 283
QY 783 rProValValSerGluHisArgLys-----GlnIleAspAspLeuLysLysPheLysAs 801
Db 282 CCCTGGTCTTCAGATGAACAGAAAGATTCCTCAACTGGAAGAACTGAGAAAGTTTGGGCG 223
QY 801 nAspPheArgLeuGlnProSerSerThrSerGlu-----SerMetAspGlnLeuLeuAsnLys 820
Db 222 CCAGTTTAAAGCTTCAGCCAGTACGCTCCCTGAGAACAGCCTGGATCTCTTCTCCCTCCCG 163
QY 820 sAsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLys 840
Db 162 G-----ATCTTAAAGGAG-----GAGCCCAAGGAAA 136
QY 840 sAspSerPheIleGluAsn-----SerSerSerAs 850
Db 135 GGAGAAAGAGTTGATGGTCTGTGACTTCAGAGCCCATGGGGTCTCCCGTCTCTCCCTCAA 76
QY 850 nCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAs 870
Db 75 GACAGAGTCCGTATCGATGAAGAGGAGACAAACCACCCCTGGCACCATCA-----GGAGG 22
QY 870 nThrGluHisLysArgProGlu 878
Db 21 CACTGAG-----GGGCCAGAG 6

RESULT 12

US-10-956-157-3976/c
; Sequence 3976, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3976
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-3976

Alignment Scores:
Pred. No.: 3,17e-46 Length: 682
Score: 917.00 Matches: 170
Percent Similarity: 90.53% Conservative: 2
Best Local Similarity: 89.47% Mismatches: 0
Query Match: 13.17% Indels: 18
DB: Gaps: 1

US-10-802-228-2 (1-1312) x US-10-956-157-3976 (1-682)

QY 1058 GlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAla 1077
Db 682 AAGTCTCAGCATCCTCATGCTATAGTCTGTAAATACAGGGTAATGCTAGAATGATGGCA 623
QY 1078 ProProThrHisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHis 1097

Db 622 CCACCAACACAGCCCGAGCTGTGTAGTATCTTCTCAGCAACTCAGTACGGGCTCAT 563
QY 1098 GluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerPro 1117
Db 562 GAGCAGACGATCGCATGTAT----- 542
QY 1118 SerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAla 1137
Db 541 -----GTTTCCACGGGCTCCCTTGTCTCAGCAGTATGCGCACCCCTAACGCT 497
QY 1138 ThrLeuHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGln 1157
Db 496 ACCTGACCCACATATCTCCACACCTCAGCCTTACCTACCCCACTCGACAGCAGCAA 437
QY 1158 SerGlnHisGlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAla 1177
Db 436 AGCCAAACATGGTGAAGTCTCTGTCAGCCAGTCTGTTCAGCACCATCAGCACCAAGGCC 377
QY 1178 AlaGlnAlaLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeu 1197
Db 376 GCCCAGGCTCTCCATCTGGCCAGTCCACAGCAGCAGTACGCCATTTTACCACCGGGGCTT 317
QY 1198 AlaProThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPhe 1217
Db 316 GGGCCAACTCCACCTCCATGACCTGCTCAACACGACGAGTCCGCCACAGAAATAGTTTC 257
QY 1218 ProAlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThr 1237
Db 256 CCAGCAGCAGCAACAGACAGTCTCTTACGATCCATCTCTTCTCAGCTTCAGCCGCGGTATACC 197
QY 1238 AsnProHisMetAlaHisValProGln 1247
Db 196 AACCCACCCACATGGCCCGCCAGCTACCTCAG 167

RESULT 13

US-10-956-157-9211
; Sequence 9211, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-9211

Alignment Scores:
Pred. No.: 3,17e-46 Length: 682
Score: 917.00 Matches: 170
Percent Similarity: 90.53% Conservative: 2
Best Local Similarity: 89.47% Mismatches: 0
Query Match: 13.17% Indels: 18
DB: Gaps: 1

US-10-802-228-2 (1-1312) x US-10-956-157-9211 (1-682)

QY 1058 GlnSerGlnHisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAla 1077
Db 1 AAGTCTCAGCATCCTCATGCTATAGTCTGTAAATACAGGGTAATGCTAGAATGATGGCA 60
QY 1078 ProProThrHisAlaGlnProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHis 1097
Db 61 CCACCAACACAGCCCGAGCTGTGTAGTATCTTCTCAGCAACTCAGTACGGGCTCAT 120

```
QY 1098 GluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerPro 1117
Db 121 GAGCAGACGCATGCGATGAT----- 141
QY 1118 SerPheTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAla 1137
Db 142 -----GTTTCCACGGGCTCCCTTGTCTCAGCAGTATGCCACCCCTAACGCT 186
QY 1138 ThrLeuHisProHisThrProHisProGlnProSerAlaThrProThrGlyGlnGln 1157
Db 187 ACCCTGCACCCACATACTCCACACCTCAGCCTTCCAGCTACCCCTCCTGGACAGAGCAA 246
QY 1158 SerGlnHisGlySerHisProAlaProSerProValGlnHisHisGlnHisGlnAla 1177
Db 247 AGCCCAACATGGTGGAGTCACTCTGCACCCAGTCTCTTTCAGCACCATCAGCACCGGCC 306
QY 1178 AlaGlnAlaLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeu 1197
Db 307 GCCCAGGCTCTCCATCTGCCAGTCCACAGCAGCAGTCCAGCAATTTACACGCGGGGCTT 366
QY 1198 AlaProThrProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPhe 1217
Db 367 GGGCCAACTCCACCTCCATGACACTGCTCCACACGCGAGTCCGCCACAGATAGTTTC 426
QY 1218 ProAlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThr 1237
Db 427 CCAGCAGCACACACACTGTCTTTACGATCCATCTCTTCAGTCCAGTCCAGCGCGGTATACC 486
QY 1238 AsnProHisMetAlaHisValProGln 1247
Db 487 AACCCACCCCATGGCCACGCTACCTCAG 516
RESULT 14
US-10-750-323-1
; Sequence 1, Application US/10750323
; Publication No. US20050032083A1
; GENERAL INFORMATION:
; APPLICANT: Pulst, Stefan M
; TITLE OF INVENTION: NUCLEIC ACID ENCODING SPINOCEREBILLAR
; ATAXIA-2 AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Mueeting, Raasch & Gebhardt, P.A.
; STREET: 119 North Fourth Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: USA
; ZIP: 55401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/750,323
; FILING DATE: 30-Dec-2003
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/727,084
; FILING DATE: 08-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueeting, Ann M.
; REGISTRATION NUMBER: 33,977
; REFERENCE/DOCKET NUMBER: 232.00010101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612/305-1220
; TELEFAX: 612/305-1228
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 516 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
```

```
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-750-323-1
Alignment Scores:
Pred. No.: 9,61e-43 Length: 516
Score: 857.00 Matches: 166
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.31% Indels: 0
DB: 8 Gaps: 0
US-10-802-228-2 (1-1312) x US-10-750-323-1 (1-516)
QY 78 GlySerAsnGlyAsnGlyGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeu 97
Db 3 GGTAGCAACGGAACAGCGCGGGCGCGGTTTCGGGCCGGCTCCCGGGCGCTCTTGGTCTC 62
QY 98 GlyGlyProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaPro 117
Db 63 GCGGGGCTCCCGCGCTTTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122
QY 118 ProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerGly 137
Db 123 CCGGCCCGCGCAACCGCGCTCCCGCTCGCGCGCGCTCGCTCCCGCGCGCTTCCCGGC 182
QY 138 ValSerLeuAlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrGly 157
Db 183 GTCTCTTGGCGCGCGCGCTCCCGGCTGTCGCCCGCGCGCTCGAGCGCGGTGTATGGG 242
QY 158 ProLeuThrMetSerLeuLysProGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 177
Db 243 CCCCTCACCATGTCTGAAGCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 302
QY 178 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 197
Db 303 CAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 362
QY 198 ProGlyGlySerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSer 217
Db 363 CCGCGCGCAGCGCGCTTCTAGCGTCCCGCGCGCGCGCTTCCCGCTCTCTCTCTCTCG 422
QY 218 ValSerSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyGly 237
Db 423 GTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 482
QY 238 ArgProGlyLeuGlyArg 243
Db 483 AGCGCCCGCGCTGGGCAGG 500
RESULT 15
US-09-918-995-28268
; Sequence 28268, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28268
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(485)
; OTHER INFORMATION: n = A,T,C or G
```

US-09-918-995-28268

Alignment Scores:
 Pred. No.: 2,74e-39 Length: 485
 Score: 799.00 Matches: 145
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 11.48% Indels: 0
 DB: 3 Gaps: 0

US-10-802-228-2 (1-1312) x US-09-918-995-28268 (1-485)

QY	1128	LeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGln	1147
DB	51	CTTGCTCAGCAGTATCGCACCCCTAACGCTACCTGCACCCACATACCTCCACACCCCTCAG	110
QY	1148	ProSerAlaThrProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaPro	1167
DB	111	CCTTCAGCTACCCCACTGGACAGCAGCAAGCCACATGGTGGAAAGTCTCTGCACCC	170
QY	1168	SerProValGlnHisGlnHisGlnAlaAlaGlnAlaLeuHisLeuAlaSerProGln	1187
DB	171	AGTCCTGTTTCAGCACCATCAGCACACGAGCGGCCAGGCTCTCCATCTGGCCAGTCCACAG	230
QY	1188	GlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProThrProSerMetThrProAla	1207
DB	231	CAGCAGTCAGCCATTTACCACGCGGGGCTTGGCCCAACTCCACCTCCATGACACCTGCC	290
QY	1208	SerAsnThrGlnSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIle	1227
DB	291	TCCAACACGCGAGTCGCCACAGAAATAGTTTCCACAGCAGCAACAGACTGTCTTTACGATC	350
QY	1228	HisProSerHisValGlnProAlaTyrThrAsnProHisMetAlaHisValProGln	1247
DB	351	CATCCTTCTCACGTTTCAGCGGGGTATACCAACCCACCATGGCCACGTACCTCAG	410
QY	1248	AlaHisValGlnSerGlyMetValProSerHisProThrAlaHisAlaProMetMetLeu	1267
DB	411	GCTCATGTACAGTCAGGAATGGTTCTTCTCATCCAACCTGCCCATGGCCAAATGATGCTA	470
QY	1268	MetThrThrGlnPro	1272
DB	471	ATGACGACACAGCCN	485

Search completed: December 23, 2005, 14:14:21
 Job time : 2122 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 23, 2005, 06:53:47 ; Search time 320 Seconds
(without alignments)
2127.715 Million cell updates/sec

Title: US-10-802-228-2
Perfect score: 6961
Sequence: 1 MRSAAAPSPAVATESRRF.....TAHFPMTHPSVQAHQQOL 1312

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4168288 seqs, 259477437 residues

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DSV=xlp
-Q=/cgn2/USFIO.spool_p/US10802228/runat_21122005_110226_9796/app_query.fasta_1.1479
-DB=Published Applications NA.New -QFMT=fastcap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10802228.ecgn 1.184 @runat_21122005_110226_9796
-NCPU=6 -ICPU=3 -NO_MAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=130 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.New:

- 1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US03_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:
- 9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:
- 10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	553.5	8.0	116856	7 US-11-143-980-1	Sequence 1, Appli
C 2	461.5	6.6	5682	7 US-11-059-982-2	Sequence 2, Appli
C 3	440	6.3	15720	7 US-11-108-172-1058	Sequence 1058, Ap
C 4	437	6.3	10968	7 US-11-075-185-35	Sequence 35, Appli
C 5	437	6.3	78869	7 US-11-075-185-1	Sequence 1, Appli
C 6	435	6.2	6210	7 US-11-056-470-1	Sequence 1, Appli
C 7	416	6.0	10524	7 US-11-075-185-37	Sequence 37, Appli
C 8	407	5.8	164810	7 US-11-121-086-4	Sequence 4, Appli

C 9	399	5.7	11070	7 US-11-075-185-34	Sequence 34, Appli
C 10	395.5	5.7	153376	7 US-11-121-086-5	Sequence 5, Appli
C 11	395.5	5.7	172543	7 US-11-121-086-6	Sequence 6, Appli
C 12	394.5	5.7	7006	6 US-10-821-234-218	Sequence 218, App
C 13	386	5.5	8730	7 US-11-087-100-1	Sequence 1, Appli
C 14	386	5.5	8730	7 US-11-087-084-1	Sequence 1, Appli
C 15	386	5.5	8730	7 US-11-087-085-1	Sequence 1, Appli
C 16	385	5.5	6360	7 US-11-056-470-2	Sequence 2, Appli
C 17	375	5.4	3408	6 US-10-858-730-40	Sequence 40, Appli
C 18	374	5.4	6683	6 US-10-995-561-473	Sequence 473, App
C 19	374	5.4	6700	6 US-10-995-561-471	Sequence 471, App
C 20	374	5.4	6833	6 US-10-995-561-474	Sequence 474, App
C 21	374	5.4	6871	6 US-10-995-561-474	Sequence 49, Appli
C 22	370	5.3	6941	6 US-10-432-483-49	Sequence 49, Appli
C 23	364.5	5.2	4598	6 US-10-955-054-71	Sequence 71, Appli
C 24	361.5	5.2	5679	7 US-11-075-185-36	Sequence 36, Appli
C 25	361.5	5.2	6012	6 US-10-467-657-83	Sequence 83, Appli
C 26	361.5	5.2	6012	6 US-10-467-657-6321	Sequence 6321, Ap
C 27	359	5.2	6763	7 US-11-080-991-55	Sequence 55, Appli
C 28	357	5.1	11736	7 US-11-000-463-218	Sequence 218, App
C 29	356.5	5.1	8651	6 US-10-432-483-48	Sequence 48, Appli
C 30	355	5.1	116856	7 US-11-143-980-1	Sequence 1, Appli
C 31	352	5.1	8651	6 US-10-432-483-48	Sequence 48, Appli
C 32	348.5	5.0	4509	7 US-11-087-100-5	Sequence 5, Appli
C 33	348.5	5.0	4509	7 US-11-087-084-5	Sequence 5, Appli
C 34	348.5	5.0	4509	7 US-11-087-085-5	Sequence 5, Appli
C 35	348.5	5.0	6786	7 US-11-069-834-59	Sequence 59, Appli
C 36	347.5	5.0	63984	7 US-11-121-086-26	Sequence 26, Appli
C 37	346.5	5.0	23907	7 US-11-186-731-6	Sequence 6, Appli
C 38	346.5	5.0	24120	7 US-11-186-731-4	Sequence 4, Appli
C 39	341.5	4.9	13650	6 US-10-995-561-155	Sequence 155, App
C 40	341.5	4.9	13711	6 US-10-995-561-154	Sequence 154, App
C 41	341.5	4.9	14194	6 US-10-995-561-153	Sequence 153, App
C 42	340	4.9	14172	7 US-11-075-185-2	Sequence 2, Appli
C 43	339.5	4.9	5913	6 US-10-821-234-789	Sequence 789, App
C 44	339.5	4.9	6360	7 US-11-056-470-2	Sequence 2, Appli
C 45	339.5	4.9	6377	7 US-11-069-834-57	Sequence 57, Appli

ALIGNMENTS

RESULT 1
US-11-143-980-1/c
; Sequence 1, Application US/11143980
; Publication No. US20050272133A1
; GENERAL INFORMATION:
; APPLICANT: He, Min
; APPLICANT: Hucul, John
; APPLICANT: Haltli, Bradley A.
; APPLICANT: Wagenaar, Melissa M.
; APPLICANT: Graziani, Edmund
; APPLICANT: Summers, Mia
; APPLICANT: Kulowski, Kerry
; APPLICANT: Pong, Kevin
; TITLE OF INVENTION: Biosynthetic Gene Cluster for the Production of a Complex
; FILE OF INVENTION: Polyketide
; FILE REFERENCE: AM-101426US
; CURRENT APPLICATION NUMBER: US/11/143,980
; CURRENT FILING DATE: 2005-06-03
; PRIOR APPLICATION NUMBER: US 60/664,483
; PRIOR FILING DATE: 2005-03-23
; PRIOR APPLICATION NUMBER: US 60/576,895
; PRIOR FILING DATE: 2004-06-03
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 116856
; TYPE: DNA
; ORGANISM: Streptomyces sp.
US-11-143-980-1

Alignment Scores: 1.13e-11 Length: 116856
Pred. No.: 116856

Score: 553.50 Matches: 383
Percent Similarity: 32.45% Conservative: 169
Best Local Similarity: 22.52% Mismatches: 613
Query Match: 7.95% Indels: 543
DB: 7 Gaps: 74

US-10-802-228-2 (1-1312) x US-11-143-980-1 (1-116856)

Qy 9 ArgSerProAlaValAlaThrGluSerArgArgPheAlaAlaAaArgTyrProGlyTyr 28
Db 74646 CGGAGCCAGCAGCAGCGCGATCCGGATCCGGTCTCGGTCTGGCGTAGGTGGCCAGCA--- 74590
Qy 29 ArgSerLeuGlnArgProAlaAlaArgSerGlyArgGlyGlyGlyAlaAlaProGly 48
Db 74589 -----GAGCCTGGCGCTCGATGGGTGCG-----CCAGCG 74560
Qy 49 ProTyrProSerAlaAlaProPro----- 57
Db 74559 TCGTACCGTACCGTGGCCCTCCACCGTCCACGTCTACTGTGGACAGAGTGGTTCG 74500
Qy 58 ProGlyProGlyProProProSerArgGlnSerProProSerAlaSerAspCysPhe 77
Db 74499 CCAGCGCTGCC-----GGATCACCGGTCTGTCGACGGAC----- 74464
Qy 78 GlySerAsnGlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyLeu 97
Db 74463 CATTCGGCGCGCTCAACCATTCGACGACCATCTCTGTTGACGGCG----- 74416
Qy 98 GlyGlyProProArgProPheValValLeuLeuProLeuAlaSerProGlyAlaPro 117
Db 74415 -----AGCCAGCATCATCCCCCGACACC 74392
Qy 118 -----ProAlaAlaProThr-----ArgAlaSer 125
Db 74391 GGTGGCGCTTCGCTCGCGTGGAGCGCGTCCACAGCAGATGCTACGCTTCGC 74332
Qy 126 ProLeuGlyAlaAlaSerProProArgSerGlyValSerLeuAlaArgProAlaPro 145
Db 74331 CCCAGCGGTGGCTGCGCGCGCGCGCGGAGC-----CCTTGCACC 74290
Qy 146 GlyCysProArgProAla-CysGluProValTyrGlyProLeu----- 159
Db 74289 GGCCATCAGCGGCCAACGGCGCTTCGCG-----GGAGAACTCGCGGAAACGGGTGGCG 74236
Qy 160 -----ThrMetSerLeuLysProGlnGln 168
Db 74235 TGGCCATTACGTGGCGCGCGCCGAGCGGCATACGGCAATTGCGCGAGCGCGCT 74176
Qy 168 nGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPr 188
Db 74175 GCGCGCCAGGTGCAATGGCCACAGCAGCAGCAACAGCGGTGTCCACGGTATGG-CC 74117
Qy 188 oProProAlaAlaAlaAsnValArgLysProGlyGlySerGlyLeuLeuAlaSerProAl 208
Db 74116 GGGCCTTCAGCCCGAAGGTGAGGACAGTGGCGCG-GATGCCACGCTCGCGCGCTGCC 74058
Qy 208 aAlaAlaProSerProSerSerSerValSerSerSerAlaThrAlaProSerSe 228
Db 74057 GGTGGCCAGATGGCGTCCGACAGTGTGGCGGTACCGAGGGCGGGTAACTCTG 73998
Qy 228 rValValAlaAlaThrSerGlyGlyArgProGlyLeuGlyArgGlyAsnSerAs 248
Db 73997 GCGGTGGTGGCGAGAAC-----ACTCGGTGGCGTGTCCGCAACGCGCTC 73950
Qy 248 nLysGlyLeuProGlnSerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMet 268
Db 73949 CCGCGCGATCCCG-----GCCCGCTCGAGACCTC 73920
Qy 268 lHisIleLeuThrSer-----ValValGlySerLysCysG 280
Db 73919 CCACGAGGCTCGACGACGAGCGGTGTGGATCCATCGGATCCGCTCGCGCGGGA 73860
Qy 280 uValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCy 300

Db 73859 GATCCCGAAGAACCGGGTTCGAATCGCGCCGCTGTAGAGGAAGCCACCGCCGAGC 73800
Qy 300 sAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSerSerGlyProLysAr 320
Db 73799 CGAGCG-CCCATCGACTCCGCCAC----- 73776
Qy 320 gGluGluIleMetGluSerIleLeuPheLysCysSerAspPheValValGlnPheLys 340
Db 73775 ----GAGTCCCGCAGCGCGGTCTGTGAAAGGCGGAAATGGCATCCGTCGCCGACTCCA 73720
Qy 340 sAspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAla-- 359
Db 73719 GGAGCCGCCACAGCTCTTCGGGGCGAGTCGACACCACGAGGAATCGGCAGCTCATGCGGA 73660
Qy 360 -----LysValAsnGlyGluHisLysGlu----- 367
Db 73659 TGATCGCAATCGGGTGTCTCATCGACCGGAGCGGTGGGACCGTCTCGAGAGTGTGCGG 73600
Qy 368 -----LysAspLeuGluProTyrAspAlaGlyGluLeuThrAlaAsnGluGluLe 384
Db 73599 ACCCTGCCCGCAGCAGTGTCTCCCGCAGGTGGCGCGGAGAAGTTCGGAGTGGGGTAGT 73540
Qy 384 uGlu-----AlaLeuGluAs 389
Db 73539 CGAAGACGACCGAGCGAGGTGTCTGCATCCCGTGGCGCCATGAGCGGATTCGCGAGTT 73480
Qy 389 nAspValSerAsnGlyTyrAspProAsn-----AspMetPh 401
Db 73479 CGATCGGGTCCAGGATCGGAAACCGGCTGACTGGAATGTCGTCGGTGTGATCGGCC 73420
Qy 401 eArgTyrAsn---GluGluAsnTyrGlyValValSerThrTyrAspSerSerLeuSe 420
Db 73419 CCGGGTGGCATGTCTCCAGAACGACGCGCACCTGGCCAGCTACGAG-----CTCAAGGA 73366
Qy 420 rTyrThrValProLeuGluArgAspAsn-----SerGluGluPheLeuLysAr 436
Db 73365 CCGTCCGTTCCTCGCTCGCGCGGGGAGGTCCCTAGCCGTCGCAACCGCCGACCG 73306
Qy 436 gGluAlaArgAlaAsnGlnLeuAlaGluIleGluSerSerAlaGlnTyrLysAlaAr 456
Db 73305 CGG-GCGGGATAGCGGGCGCCACCAGCAATTCACCGACGAGGGGTGGCGCGACG 73247
Qy 456 gValAlaLeuGluAsnAspAspArgSerGluGluGlySerThrAlaValGlnArgAs 476
Db 73246 G-----CGGTGAACCGGGAGCGCAAGCTC 73223
Qy 476 nSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProProGl 496
Db 73222 GACCAATCCACGTCGCGCATCATCAGTCGGTCTCTGTCGTCTCAACGCTCTGTAAGG 73163
Qy 496 yGlnArgAsnArgGluValIleSerTyrGlySerGlyArgGlnAsn-----Se 512
Db 73162 GCTGCCAAGCCAGCGGTGGTGGTCCATGSCCGTGGTGGCCAGCCCGCATGCGCTGTGCG 73103
Qy 512 rProArgMetGlyGlnProGlySer-----GlySerMetPr 524
Db 73102 GCCCAGGGTCCCCAAGCCACCGACGTCGSCCACCACCCCTCCGCCGCGCCCAAGCC 73043
Qy 524 oSerArgSerThrSerHisThrSerAspPheAsn-----ProAsnSe 538
Db 73042 AACCGTCCCAACCGCATTCGCGCGCGCATAACTCCCTTCGACCGCATACCCACAGATC 72983
Qy 538 rGlySerAspGlnArgValValAsnGlyGlyValProTyrProSerProCysProSerPr 558
Db 72982 CCCGAGCCGGAAGAAACACACAA-----ACGCATCCAGATCCCAACCC 72938
Qy 558 oSerSerArg-----ProProSerArgTyrGlnSe 568
Db 72937 CGCGTCAACTCATCCAGATACCGCGCACCTCCGCTTGACCCGATCACCTCAGCAACC 72878
Qy 568 rGlyProAsnSerLeuProProArgAlaAlaThrProThrArgProProSerArgProPr 588

Db 72877 CGCTCAGAGTCCAGCGACTCCAAACACACCGTTCATCCAGACACACCGCGCGCATGCACACC 72818
Qy 588 o-----SerArgProSerArgProSerHisProSer-----AlaHisG1 602
Db 72817 ACACGCAATCCCAATCCCTCCAACTCCGACACCGCCGACCGCCGATCCGCGCACATCA 72758
Qy 602 ySerProAlaProValSerThrMetPro-----LysArgMetSerSerGluG1 618
Db 72757 CACGCGACACCGCCGACCGCAACCCCAACCCCTCCAACTCCGCGCACCGACCTGCTCCGCA 72698
Qy 618 yProProArgMetSerPro-----LysAlaGlnArgHisProArgAsnHisArgVa 635
Db 72697 TCCGCGCAGCATCCCGCGACGACTCGCGCAGCACCAAGTCTCCACACCGACACCGCGCC 72638
Qy 635 lSerAlaGly-----ArgGlySerIleSerSerGlyLe 646
Db 72637 AACCCCGCGCCACATGACACCCACGACCGGTACCGCGCTCACCAACCGCGCCCT 72578
Qy 646 uGluPheValSerHisAsnProSerGluAlaAlaThr-ProProValAlaArgThrS 666
Db 72577 CGACC--CGGCCACACACCGCGCCACCGCATCCCGAACCCGAGCCGACACCGCGCACCA 72520
Qy 666 ex-----ProSerGlyGlyThrTrpSerSerValVal----- 676
Db 72519 GCGGAGCACCACCGCCACACCAACCGAAACCGAATCTGGTCTCTCCCGGTACCCCCAC 72460
Qy 676 ----- 676
Db 72459 CCNACACCGCGCACACCAAGCCCGCTCTCATCCACCACCGCGCACATCGATCA 72400
Qy 677 -----SerGlyValProArg-----LeuSerProLysThrHisArgProA 690
Db 72399 GACCAACCCCAACGATCCGATGCTCCAAACGGATCACCGCGCCAGACCCACACCGCCG 72340
Qy 690 rGSerProArgGlnAsn----- 695
Db 72339 ACTGCCCCCGATCCACCATCATCATCCAGCACCACCGCACCGCGAGTCACACACC 72280
Qy 696 -----SerI 697
Db 72279 ACNACCGCGCATCGGCGCGCGCTCCCGAGTCCCTGCACCATGACACCGTCCGCTCGA 72220
Qy 697 leGlyAsnThrProSerGlyProValLeuAla---SerProGlnAlaGlyIleIleProT 716
Db 72219 CCGGCAACACGACACACACCGCGACCGCGACCTCGCCAGCTCGCGCATGTGCACAA 72160
Qy 716 hrGluAlaValAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSerA 736
Db 72159 CTGCGCATCGGCGCGCGCGACCGCGACGACGCGGACCATGCTCTCCGCGCTCGCCCA 72100
Qy 736 snArgAlaValThrPro-----SerSerGluAla----- 745
Db 72099 CCACGCGCATCGCGCGAGACTCGGCTCGGAGCGGTCCGATCGGGGACCATG 72040
Qy 746 -----LysAspSer-----ArgLeuGlnAspG 753
Db 72039 CGAGGCGGTACCGCGGATCTCGACGAGTACCTCTCTGTCGACGCGCGCGCGACG 71980
Qy 753 lnArgGlnAsnSerProAla-----GlyAsnLysGluAsnIleLysProAsnGluThrS 771
Db 71979 AGAGGCGCGCGAGTGGCGCGCTGAGCGGCTGTTCGGCATCGATACCGAAAGACCGAGAT 71920
Qy 771 erProSerPheSerLysAlaGluAsnLysGlyIleSerProValVal----- 786
Db 71919 CCCACGCTCAACCGCATCCAGAACCCCGGTCGCGCGCTCGTGTGGTGTGCTCAG 71860
Qy 787 -----SerGluHisArgLysGlnIleAspLeuLysPheLysAsnAspP 803
Db 71859 GGGACGCTCCAAACGACGCTCAGC-CTGGAAGGCATACGTTCG----- 71814
Qy 803 heArgLeuGlnProSerSerThrSerGlu-----SerMetAsp----- 815
Db 71813 --CAAAATCAACCGCGCGACCGCGGAAACACCGCATCTCCAGTCCGACACCGCGG 71756

Qy 816 -----GlnLeuLeuAsnLysAsnArgGluGlyGluLysSerArgAspLeu-----lleLysA 833
Db 71755 ACGTACAGCCCAACCAAGCGCCATAACCGGTTCAGAGCTCCGAGCATCCCGACGCAAG 71696
Qy 833 spLys--lle-GluProSerAlaLysAspSerPheIleGluAsnSerSerSerAsnCys 851
Db 71695 ACAGGACACACATCCCATTCGCGAAG-----CGGTCAACGTC 71657
Qy 852 ThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnThr 871
Db 71656 CGGTACGCGCCCAACTCCAGGAACCAACCAACCGAGTTCACGAGCGTCTCCAGCCCA 71597
Qy 872 GluHisLysArgGlyProGluValThr-----SerGlnGlyValGlnThrSer 887
Db 71596 TCCGGAATTCGAGCGTCTCCGCAATGACGACCCCAATATCTCCGCGAGCAGAGTCT 71537
Qy 888 SerProAlaCysLysGlnGluLysAspAspLysGluGluLysAspAlaAlaGluGln 907
Db 71536 TCACCGCGCCACCGACCGGACAGTTCGACACCGGACCGCTCGGCGACCGGAATCC 71477
Qy 908 ValArgLysSerThrLeuAsnProAsnAlaLysGluPhe----- 920
Db 71476 ACGCCCGGAGCACCC--TCACCGAATCTCCGCGAGCATCCCGTTCATACGAGCCGAATGA 71420
Qy 921 -----AsnProArgSerPheSerGlnProLysProSerThrThr-ProTh 935
Db 71419 AACGCGATGACTCACCGCAACCGCGGTCTCCGACCCCAACCGCCACCAACATCCACA 71360
Qy 935 rSerProArgProGlnAlaGlnPro-----SerProSerMetValGlyHisGlnGlnPr 953
Db 71359 ACGGCTCAACCGCATCTCTATCCCGGAGACACCAACCGATCGGAGCATTCACCGCC 71300
Qy 953 oThrProValThrGlnProValCysPheAlaProAsnMetMetTy-ProValProVa 973
Db 71299 GCNACCAACACCTCCGCGAGCGCGCGACCTCACCTCCGACCGCGCCACCGCA 71240
Qy 973 l-----SerProGlyValGlnProLeuTyProIleProMetThrProMetProVa 990
Db 71239 ACCATCGCACCCCGACCGCAACGCTGCATCAACGCGCGCGCGCCACCGACCGCTC 71180
Qy 990 lAsnGlnAlaLysThrTyArgAlaValProAsnMetProGlnGlnArgGlnAspGlnHi 1010
Db 71179 CCGCATCTCCAAACGACACACACCGCGCATATCGCGCAGCCCAACT----- 71131
Qy 1010 shHisGlnSer-----AlaMetMetHisPro-----AlaSe 1020
Db 71130 -CACCAATCGAATGACCCAGCAGATAATCCCGCTCACACCCCGAGGCGATCAGCGCA 71072
Qy 1020 rAlaAlaGlyProProIleAlaAlaThrProAla----- 1032
Db 71071 TACAGCGACACCTCCAGCGGAAACCGCGCTGCGCATACACCGTCTCCCGACCAAC 71012
Qy 1033 -----Ty-SerThrGlnTyValAlaTy-SerProGlnGlnPheProAsnGlnProL 1050
Db 71011 CCGCATCATACCCCGACACCTCCCGCGCGCTATCCAGATACCGATCCAGTCC 70952
Qy 1050 euValGlnHisValProHisTyGlnSerGlnHisProHisValTySerProValIleG 1070
Db 70951 CGGCACACCTCATCCACAC--CGCAGCGCAACAC-----CGGATACCTCTCATAC 70904
Qy 1070 lndGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerS 1090
Db 70903 AA-----CCACGACCCCATCCCAACCGCT--CGACCCCTTGACC 70866
Qy 1090 erAlaThrGlnTyGlyAlaHisGluGln-----ThrH 1101
Db 70865 CGCAAAACACACACCAACCCACCGCAACCGGAACACCCACCGACCGCAACCCCAACCC 70806
Qy 1101 isAlaMetTyAlaCysProLysLeuProTyAsnLysGluThrSerSerPheTyP 1121
Db 70805 AC-----CCAGCAACTCTCTCCAGTCCGACCGACCGACCGAC----- 70765

Db 6505 CCACGAGCCCGCGCGCTCGGTGTGCCGAGGTTGGACTTGATCGAGCGAGCCACGA 6446
Qy 199 -----GlyGlySerGly-----Leu-LeuA 205
Db 6445 GCGGCGCTCGCGGTTGTGCGGTTGCCGTAGTCCGACAGCGCTCGCGCTCGATGG 6386
Qy 205 laSerProAlaAlaProSerProSerSerSerValSerSerSerAlaThrA 225
Db 6385 GATCGCCGAGGCGCGTCAGTGCCTGCGCTCGACGCGCTCGATGTCCGCGGGTCA 6326
Qy 225 laProSerSerValAla-----231
Db 6325 GCGCGCGCGCGCGCTCGCGGATGACTCTCTGTGGCGCGCGTGTGGGCGG 6266
Qy 232 -----AlaThr-SerGlyGlyArgProGlyLeu 241
Db 6265 TGAGCCCTGGCTGCGGCGCTCTGTTGACGCGGAGCGCGGAGGACGCGCACGCG 6206
Qy 242 GlyArgGlyArgAsnSerAsn-----LysGlyLeuProGlnSerThr 255
Db 6205 GGTGCGCTGCGCGCGCGCTCGACGCGCTTGAGCACACGAGC-CCGACGCGCTCC 6147
Qy 256 IleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLeuThrSerVal 275
Db 6146 GCGGGCGGATCGGTTCCCGCGCGCGAGGCTTGACCGTGTCCGGGCGCCATG 6087
Qy 276 GlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluGlyValPheLysThr 295
Db 6087 -----6087
Qy 296 TyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSer 315
Db 6086 -----CCGCGCTGTGGGTGAA-CTCGATGAAGGCCATGGCGTGGCCATCACGGTGGC 6034
Qy 316 SerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAspPheVal 335
Db 6033 GCGCGCGCGAGGCGAGTGCCTCTCGCGCGCG-----5998
Qy 336 ValValGlnPheLysAspMetAspSerSerTyrAlaLysArgAspAlaPheThrAspSer 355
Db 5997 -----GAGCGCATGTCAGCGGAGTGGAGGACACAGCGCA 5962
Qy 356 AlaIleSerAlaLysValAsnGlyGlu-----HisLysGluLysAsp 369
Db 5961 CGACGACGACGCGCTGTTCAGCGGTACCGCGCGCGCGCGAGCCCGCGTGTAGCGAT 5902
Qy 370 LeuGluProTrpAspAlaGlyGluLeu-----ThrAlaAsnGluLeuGlu 385
Db 5901 GCGCGCGGAGCGCTGGGAAGCTGCGGTGGCGAGCTGAGCGTGAAGACCTCGAG 5842
Qy 386 AlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMetPheArgTyrAsnGlu 405
Db 5841 GTGCTCGAGCGCGCGCG-----5821
Qy 406 GluAsnTyrGlyValValSerThrTyrAspSerSerLeuSerSerTyrThrValProLeu 425
Db 5820 -----GTAGTCGCTGTACACCGAGCCGCGAGAGCGCGGTGGCGCTCGCCTC 5773
Qy 426 GluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla-----AsnGlnLeuAla 444
Db 5772 GAGCGA-----ATGGGGCGGAGGCGCGCGCTCGAGCGCTCGCCTC 5731
Qy 445 GluGluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGluAsnAspArg 464
Db 5730 GCGCACTCGACGACGCGGTGTGCGGGTCCATGCGCTCCGCTTCGCGGCGCTGAT 5671
Qy 465 SerGluGluGluLysTyrThrAlaValGlnArgAsnSerSerGluArgGluGly-----482
Db 5670 GCCGAAGAAGTTCGATCGAAGCGGT-CAGCGTCTAGAGGAAGCGCGCGCTGGTGA 5612
Qy 483 -----HisIleAsnThr 487
Db 5611 CGACTTCCCGGGGCGTCCGGTTCGGGGTGTACAGCCCGCCACATCCCGAGCGCGCC 5552

Qy 488 -----ArgGluAsnLysTyrIleProPro 495
Db 5551 CCTCGGGGAACGGCCCGATGGCTCCCGCCCTCCGAGAGCAGCTGCCACACGCGCGCG 5492
Qy 496 GlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMet 515
Db 5491 GGTATCACGCGCGCGCGAGCGCGCATCGACACATCGCATCGCTCGCTCGCT 5432
Qy 516 GlyGlnPro-----GlySerGlySerMetProSerArgSerThr 528
Db 5431 GCGCTCCCGCGCGCGCGCTGTGCGCGCTGCCCGCGCGCGCTCGAGCTCCGAGA 5372
Qy 529 SerHisThrSerAspPheAsnProAsnSer-----Gly 539
Db 5371 ACGCAGCTGAAGCAGCAGATCCCGCATCGCGCGCGCTCGGATGGTTCGAACGCGAGG 5312
Qy 540 Ser-----AspGlnArgValValAsnGlyGlyValProTrpProSerPro 554
Db 5311 TCGCGGCAACGCGCTCTCGCGCGCGCGCGAGCGGTTCCTCAGCTCGACGCGCTGA 5252
Qy 555 -----CysProSerProSerArgProProSerArgTyrGlnSerGlyProAsn 571
Db 5251 GCGAGTGAGCCCGAGCTCTTCAGCGCTTGTCCG-----CGGGAGCTCGG 5204
Qy 572 SerLeuProProArgAlaAlaThrProThrArgProProSerArgProProSerArgPro 591
Db 5203 CGGACCTGCCAGCTGCAGCACGCGCG-----CGCTCTCGCGCGCACCGCTCGAGCA 5150
Qy 592 SerArgProProSerHisProSerAlaHisGlySerProAlaProValSerThrMetPro 611
Db 5149 GCGCGCGCGCGCTCGCGCTCGGCA-----GGCGCG-----5117
Qy 612 LysArgMetSerSerGluGlyProProArgMetSerProLysAlaGlnArgHisProArg 631
Db 5116 ---AGAGCGCGCGCGAGCGCGCTCGCTCGCGCGCGCGCGCGAGCGCGCGCG 5060
Qy 632 AsnHisArgValSerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHis 651
Db 5060 -----5060
Qy 652 AsnProProSerGluAlaAlaThrProProValAlaArgThrSerProSerGlyThr 671
Db 5059 -----GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGCA-----5024
Qy 672 TrpSerSerValSerGlyValProArgLeuSerProLysThrHisArgProArgSer 691
Db 5023 -----GCGCGCGCACCTGCGCGCGCGCTCGCTCG 4997
Qy 692 ProArgGlnAsnSer---IleGlyAsnThrProSerGlyProValLeuAlaSerProGln 710
Db 4996 CCGCAGCTGGAGCGCGCACAGGTTCGAGCGCGCGCGCACCA-----GCGCGCGCTCCG 4943
Qy 711 AlaGlyIleIleProThrGluAlaValAlaMetProIleProAlaAlaSerPro---Thr 729
Db 4942 GCGCGCGAGCGCGCGGTTCAGCAGCTCATCCCTGCGCGAGCGCGATCGCGCGCATGC 4883
Qy 730 ProAla-----SerProAlaSerAsnArgAlaValThrProSerSerGluAlaLys 746
Db 4882 CTGCGCTCTGAGGCGTCCAGGTCCGCGCGCGCGCGAGGTGCGCGCTCATGCGAGCCCGG 4823
Qy 747 AspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnIleLys 766
Db 4822 CTGCTCCAGAAAGCCCGAGGAGGTTCATCGCGCGCACCGCGCGCGCGAGCTCGG 4763
Qy 767 Pro-----AsnGluThrSerPro-SerPheSe 775
Db 4762 CCGAGCGCGTCCAGGAACGCGTTGGCGCGCGGTAGTTCCGCTCGCGCGCTGCCCA 4703
Qy 775 rLysAla-----GluAsnLysGlyIleSerProValSerGluHisAr 790
Db 4702 GGTGCGCGCGCGCGAGAGAACAGCAGCAAGCGCGAGCGGATGGCTCGCTCGCTCGCG 4643

QY 790 glysGlnIleAspAspLeuLysLysPhelYsAsnAepPheArgLeuGlnProSerSerTh 810
DB |||:|||||
4642 CGGCCAGGTGCCAGCGCC-----GTGATCTTCGGCCCGCAGCACGC 4601
QY 810 rSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyGluLysSerArgAspLe 830
DB |||:|||||
4600 GCGGAGTGTCTC-----GGCGGAGAGT 4577
QY 830 uLeLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsn-SerSerSera 850
DB |||:|||||
4576 TGTACAGATGAGTGTGTCCAGCACGCCGC-----CATGTGCACGACCGCGTGGGG 4523
QY 850 snCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSera 870
DB |||:|||||
4522 GGTGTGCACCGCGGATCGCGGATACCGCCCGCCAGCGCGCGCGTGGCG----- 4470
QY 870 snThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSer-Pro 889
DB |||:|||||
4469 --ACGTGCACCGCGGAGTGTGCACGTTCGGCGCGCGCGCGCGCGTGTCCACCA 4412
QY 890 AlaCysLysGlnGluLysAspAspLysGluGlu-----LysLysAspAlaAlaGlu 906
DB |||:|||||
4411 GCGCGCGCGTCCGCGCATCCATCCGCGCGCGAGTGACGAGTGTCCGCGCACGC 4352
QY 907 Gln-----ValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
DB |||:|||||
4351 CGTGCAGCGGACAGGTGTCCGCGCGACCTCCCGCGCGAGTCTCCCGGTGCGCGCGTGA 4292
QY 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
DB |||:|||||
4291 TCAGCACGTCCCGCGGATCCAGCGCGTGTGTGCGCGCGTGTGCGCGCGCGCGCGCG 4232
QY 941 Ala-----GlnProSerProSerMetValGlyHisGlnGlnPro 953
DB |||:|||||
4231 GCACCTCGCGAGCGTGGCGCGCGCGCGCGCGCGCGCGCGAGTTCGGTCTCT 4172
QY 954 ThrProValThrThrGlnProValCysPheAlaProAsnMetMetThrProValProVal 973
DB |||:|||||
4171 CGGCCG-----CGCCA 4160
QY 974 SerProGlyValGlnProLeuThrProLeuProMetThrPro-----MetPro 989
DB |||:|||||
4159 CGCGCGCGCGCGCGCGCGTTCGTTCAGCGCGTGGCGCGCGCGCGCGCGCGCGCGCG 4100
QY 990 ValAsnGlnAlaLysThrThrArgAlaValProAsnMetProGlnGlnArgGlnAspGln 1009
DB |||:|||||
4099 CCGCGCGCGGTTC-TCTGCGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 4059
QY 1010 HisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThr 1029
DB |||:|||||
4058 ---CACAGCGCGGTGGTCAACCGCGCGCGCGCGCGTGTGCGCGCGCGCGCGCGCGCG 4002
QY 1030 ProProAlaThrSerThrGlnThrVal-----1038
DB |||:|||||
4001 CCGCGGTCAACACCGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGC 3942
QY 1039 -----AlaThrSerProGlnGlnPheProAsn 1047
DB |||:|||||
3941 ACCAGCGACAGGACCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3882
QY 1048 GlnProLeuValGlnHisValProHisThrGln---SerGlnHisProHisValThrSer 1066
DB |||:|||||
3881 TCGCCCGCGCGGTGTGATCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3822
QY 1067 ProValleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGln-----1083
DB |||:|||||
3821 GCGATGAGTCCGCGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGC 3762
QY 1084 ---ProGlyLeuValSerSer-----1090
DB |||:|||||
3761 CCGCGGAGCGAGTGTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGC 3702
QY 1091 ---AlaThrGlnThr-GlyAlaHisGluGlnThrHisAlaMetThrAlaCysProLysLe 1109

DB |||:|||||
3701 CACGTGACCGGTACAGGT---CACGCTCGTCGACCCCGCGCGCGCTTTCGCGCGTGC 3645
QY 1109 uProThrAsnLysGluThrSerProSerPheThrPheAlaIleSerThrGlySerLeuAl 1129
DB |||:|||||
3644 TCCG---CGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGC 3588
QY 1129 aGlnGlnThrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSe 1149
DB |||:|||||
3587 GGTGCGCGCGTGTGATCGACCGCGCGAGTGTGTCCCATCGTGTGCGTCTCGACCTCC 3528
QY 1149 rAlaThrProThrGlyGlnGlnSerGlnHisGlyGlySerHisProAlaProSerPr 1169
DB |||:|||||
3527 ACGCCGTCC 3480
QY 1169 oValGlnHisHisGlnHis-----GlnAlaAlaGlnAlaLeuHisLeuAlaSerPr 1186
DB |||:|||||
3479 GACCGCGGAGGCGATCGGATGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGTCC 3420
QY 1186 oGlnGlnGlnSerAlaIleThrHis-----AlaGlyLeuAlaProThrPr 1201
DB |||:|||||
3419 CGCAGGAGCGCGTGGAGCGCGCGCATCGACAAAGCGCGGATGCAAGCGCGAGTCTCGCGCC 3360
QY 1201 oProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAla-----1219
DB |||:|||||
3359 CGC-TCGAGACCGCTCGGCAACCGCAGGTGCGGTAGATCGACGCGCGCGCGCGCGCGCG 3301
QY 1220 -AlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaThr-----1237
DB |||:|||||
3300 CGCGCGCGTACGCGCGCGGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3241
QY 1238 -----AsnProHisMetAlaHisValProGlnAlaHis-----1249
DB |||:|||||
3240 CGTGCAGTGCAGCGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 3181
QY 1250 -ValGlnSerGlyMetValProSerHisProThrAla-----His-----1262
DB |||:|||||
3180 GCGCGCGCGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGC 3121
QY 1263 -AlaProMetMetLeuMetThrThrGlnProProGlyGlyProGlnAlaAlaLeuAlaG1 1282
DB |||:|||||
3120 CGCGCGTGTGCGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGC 3061
QY 1282 nSerAlaLeuGlnProIleProValSerThrThrAla 1294
DB |||:|||||
3060 CACC-----GTCACTGTGCGCGTCAACCGCG 3036

RESULT 5

US-11-075-185-1/c
; Sequence 1, Application US/11075185
; Publication No. US20050266434A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMBRUTICINS
; FILE REFERENCE: 010099.03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 78869
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-11-075-185-1

Alignment Scores:

Pred. No.: 2.74e-07 Length: 78869

Score:	437.00	Matches:	353
Percent Similarity:	32.24%	Conservative:	149
Best Local Similarity:	22.67%	Mismatches:	592
Query Match:	6.28%	Indels:	470
DB:	7	Gaps:	66

US-10-802-228-2 (1-1312) x US-11-075-185-1 (1-78869)

QY	18	ArgArgPheAlaAlaAlaArgTrpProGlyTrpArgSerLeuGlnArgProAlaArgArg	37
DB	22402	CGTCGAGCGCGCGGAGACGGGTGGTGGGTACAGGCGCGCCCATGCGCGCGC	22343
QY	38	SerGlyArgGlyGlyGlyAlaAlaProGlyProTyProSerAlaAlaPro-----	55
DB	22342	GCTGGCTGCTCCCTGGCGGTGAAGAGACCGCACCTTG-CCGGGGGGTTCGCCCGCGCGC	22284
QY	56	-----ProProProGlyProGlyProProPro-----SerArgGlnSerPro	70
DB	22283	GTACGCGCGCGCGCTGGCGCGCGCGCTGAGCGGAACTCGGCAAGCGGCGCGCC	22234
QY	71	ProSerAlaSerAspCysPheGlySerAanGlyAsnGlyGlyGlyAlaPheArg-----	88
DB	22223	AGCTCTCGTGGCTGCGTCCGCGCGAGCGAGCGA--GCCGCGTGGCGAGGTGCG	22166
QY	89	-----ProGlySerArgArgLeuLeuGlyLeuGlyGlyPro-----	100
DB	22165	TGCGCGTCTGTCGCGAGGTTCGCGCGACGTCGAGCAGCGCGTGGTCCGGTGGCGCGCA	22106
QY	101	---ProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAla	119
DB	22105	GGTGTCTCG-----CGAGCGCGCGCGCTCGGCGC-----GCA	22073
QY	120	AlaProThrArgAlaSerProLeuGlyAlaArgAlaSer-----ProPro	134
DB	22072	GGCGCGCTGTCGCGACCGCGAGGAGCGCGGCGAGGTTCACCGCACCGCGCGCGCG	22013
QY	135	ArgSerGlyValSerLeuAlaArgProAlaProGly-----	146
DB	22012	CCTCGCGCGCGCTCTCGCGCGCGCGCGCTCTCGACGATGACGTGCGGTGGTGC	21953
QY	147	-----Cys-----ProArgProAla-----Cys-GluProValTy	156
DB	21952	CGCTGATCCGAGGACGAGACCGCGCGCGCGCGCGCGTCCGTGCGCGCGCACGGCA	21893
QY	156	rgly-----ProLeuThrMetSerLeuLysProGlnGlnGlnGln-----	170
DB	21892	CGGGCTGCTCAGGAGTTGATGCTCCCTGCGACCATGTCGATGTCGGCGATCGGGCT	21833
QY	171	-----GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPr	188
DB	21832	CGGCATACAGCGCTCTGGCAGCTCTGCGTGTGTCATGCCACGACACGATTCATCAGCC	21773
QY	188	oProProAlaAlaAlaAsnValArgLysPro-----	198
DB	21772	CCACGACGCGCGCGCGCTGCGTGTCCGAGTTGGACTTGTATCGAGCGAGCCACGA	21713
QY	199	-----GlyGlySerGly-----Leu-LeuA	205
DB	21712	GCGGCGCTCGCGGTGTCGCGGTGCGGTAGTTCGCCAGCAGCGCTGCGCTCGATGG	21653
QY	205	laSerProAlaAlaAlaProSerProSerSerSerValSerSerSerAlaThrA	225
DB	21652	GATCGCCGAGGCGCGTCCAGTCCGTCGCTCGACGCGCTGATGTCGCGCGGTC	21593
QY	225	laProSerSerValAlaAla-----	231
DB	21592	GCCCCCGCGCGCGCGCTGGCGGATGACCTCTGTCGCGCGCGCGGTGGCGCGG	21533
QY	232	-----AlaThr-SerGlyGlyArgProGlyLeu	241
DB	21532	TGAGCCCTTGGCTGCGCGCTCTGTTGACCGCGGAGCGCGGAGGACGCGCACGCG	21473
QY	242	GlyArgGlyArgAsnSerAsn-----LysGlyLeuProGlnSerThr	255

DB	21472	GGTCCCGTCCGCGCGCGCTCCGACAGCCGCTTGAGCACCGAGGC-CCGACGCGCTCC	21414
QY	256	lleSerPheAspGlyIleTyAlaAsnMetArgMetValHisIleLeuThrSerValVal	275
DB	21413	CGGGGCGCATGCGCTTCGCGCGCGCCCGAAGGCTTGACCGTCCGTCGGGGGCCATG	21354
QY	276	GlySerLysCysGluValGlnValLysAsnGlyGlyIleTyArgGlyValPheLysThr	295
DB	21354	-----	21354
QY	296	TyrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGluSerSer	315
DB	21353	-----CCGCGCTGTCGCTGAA-CTCGATGAAGCCATGGGCGTGGCCATCACGTTGC	21301
QY	316	SerGlyProLysArgGluGluLeuMetGluSerIleLeuPheLysCysSerAspPheVal	335
DB	21300	GCGCGCGGAGGCGCAGGTGCGACTCCCGCGCGG-----	21265
QY	336	ValValGlnPheLysAspMetAspSerSerTyAlaLysArgAspAlaPheThrAspSer	355
DB	21264	-----GAGCGACATGACCGCGAGGTGAGGAGCACACGAGCGA	21229
QY	356	AlaIleSerAlaLysValAsnGlyGlu-----HisLysGluLysAsp	369
DB	21228	CGACGAGCAGCGCTGTCGACGCTGACCGCGCGCGCGCGAGCCCGCAGCGTGTAGCGAT	21169
QY	370	LeuGluProTrpAspAlaGlyGluLeu-----ThrAlaAsnGluLeuGlu	385
DB	21168	GCGCCCGAGCGCGCTGGGAAAGCTCCGCTGGCGACGTAGCGTCAAGACCTCGAG	21109
QY	386	AlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMetPheArgTyArgGlu	405
DB	21108	GTGCTCGAGACCGCGCGCC-----	21088
QY	406	GluAsnTyGlyValValSerThrTyArgSerSerLeuSerSerTyThrValProLeu	425
DB	21087	-----GTAGTCGCTGACACGAGCGCGCGAGCGCGCGCTCGCGCTCGCCTC	21040
QY	426	GluArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla--AsnGlnLeuAla	444
DB	21039	GAGCGA-----ATGGGCGCGAGCGCGCGCGCTCGAGCGCTCCCA	20998
QY	445	GluGluIleGluSerAlaGlnTyLysAlaArgValAlaLeuGluAsnAspArg	464
DB	20997	GCGCGACTCGAGCAGCAGCGCTGCTGCGGTCCATCGCTCCGCTTCGCGCGGCTGAT	20938
QY	465	SerGluGluGluLysTyThrAlaValGlnArgAsnSerSerGluArgGlyGly---	482
DB	20937	GCCGAAGAAGTTCGGATCGAAGCGGT-CAGCGTCTGTAGAGGAAGCGCGCGAGTTGGTCA	20879
QY	483	-----HisSerIleAsnThr	487
DB	20878	CCGACTTCCCGGGGCGTTCGCGGTGCGGTGTCACAGCCCGCGCCACATCCCGCGCGCC	20819
QY	488	-----ArgGluAsnLysTyIlePro	495
DB	20818	CCTCGGGGAAAGCGCGCGATGGGTCCCGCGCTCCGAGAGCAGCTGCCACAGCGCGCGG	20759
QY	496	GlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMet	515
DB	20758	GGGTATCAGCGCGCGCGGAGCGCGCATCGACACGATCGCGATCGGCTCGCTCGT	20699
QY	516	GlyGlnPro-----GlySerGlySerMetProSerArgSerThr	528
DB	20698	GCGCTCCCGCGCGCGCGGTGTCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	20639
QY	529	SerHisThrSerAspPheAsnProAsnSer-----Gly	539
DB	20638	ACGACGCTGAAGCAGCAGATCCGCGATCGCGCGCGCGCTCGGATGGTTCGAACGCGAGG	20579
QY	540	Ser-----AspGlnArgValValAsnGlyGlyValProTrpProSerPro---	554

Db 20578 TCAGCGGAAACGGCTCTCGGCGCGCGCGGAGGGCGGTTCCTCAGCTCGACGGCGCTGA 20519
Qy 555 -----CysProSerProSerArgProProSerArgTyrGlnSerGlyProAsn 571
Db 20518 GCAGGTGAGCCCGCTCTTCAGCGGCTGTTCG-----CGGGACCTGCG 20471
Qy 572 SerLeuProProArgAlaAlaThrProThrArgProProSerArgProProSerArgPro 591
Db 20470 CCGGACCTGCGACCTGCGACGCGCG-----CCGCTTCGCGCGACACGAGCTCGAGCA 20417
Qy 592 SerArgProProSerHisProSerAlaHisGlySerProAlaProValSerThrMetPro 611
Db 20416 GCAGCGCGCGCGCTCGGCTCGGCA-----CGCGCG----- 20384
Qy 612 LysArgMetSerSerGluGlyProProArgMetSerProLysAlaGlnArgHisProArg 631
Db 20383 ---AGAGCGCGCGCGCGCGCTCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 20327
Qy 632 AsnHisArgValSerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValSerHis 651
Db 20327 ----- 20327
Qy 652 AsnProProSerGluAlaAlaThrProProValAlaArgThrSerProSerGlyGlyThr 671
Db 20326 -----GCCG 20291
Qy 672 TrpSerSerValValSerGlyValProArgLeuSerProLysThrHisArgProArgSer 691
Db 20290 -----GCCG 20264
Qy 692 ProArgGlnAsnSer---IleGlyAsnThrProSerGlyProValLeuAlaSerProGln 710
Db 20263 CCGCAGCTGCGAGCG 20210
Qy 711 AlaGlyIleIleProThrGluAlaValAlaMetProIleProAlaAlaSerPro---Thr 729
Db 20209 GCGCGCGAGCG 20150
Qy 730 ProAla-----SerProAlaSerAsnArgAlaValThrProSerSerGluAlaLys 746
Db 20149 CTGCGCTGTGAGCGGTGCGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 20090
Qy 747 AspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAlaLys 766
Db 20089 CTTGCTCCAGAACCCCG 20030
Qy 767 Pro-----AsnGluThrSerPro-SerPheSe 775
Db 20029 CCGGAGCGGTCCAGGAACGGTTGGCGCGCGGTAGTTGCGCTGCGCGCGCGCGCGCGCG 19970
Qy 775 rLysAla-----GluAsnLysGlyIleSerProValValSerGluHisAr 790
Db 19969 GCGTGGCGCGCGCGCGAGACAGACAGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19910
Qy 790 gLysGlnIleAspAspLeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerTh 810
Db 19909 CGGCGAGGTGCAGCGCGCG-----GTGATCTTCGCGCGCGCGCGCGCGCGCGCGCG 19868
Qy 810 rSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyLeuLysSerArgAspLe 830
Db 19867 GCGCGAGCTGCTC-----GCCGCGAGAGCT 19844
Qy 830 uilleLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsn-SerSerSerA 850
Db 19843 TCGTACGATGACGTGCTCCAGCAGCGCGCG-----CATGTGACACCGCGCGCGCGCG 19790
Qy 850 snCysThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerA 870
Db 19789 GGTGTGCGCGCGGATGCGCGCGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19737
Qy 870 snThrGluHisArgGlyProGluValThrSerGlnGlyValGlnThrSerSer-Pro 889
Db 19736 --ACGTGCGCAGCG 19679

Qy 890 AlaCysLysGlnGluLysAspAspLysGluGlu-----LysLysAspAlaAlaGlu 906
Db 19678 GCG 19619
Qy 907 Gln-----ValArgLysSerThrLeuAsnProAsnAlaLysGluPhe 920
Db 19618 CTTGCG 19559
Qy 921 AsnProArgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGln 940
Db 19558 TCAGCACCTGCTCCCGCGCGGATCCAGCGCGCTCGTCCGCGCGCGCGCGCGCGCGCG 19499
Qy 941 Ala-----GlnProSerProSerMetValGlyHisGlnGlnPro 953
Db 19498 GCACCTCG 19439
Qy 954 ThrProValTyrThrGlnProValCysPheAlaProAsnMetMetTyrProValProVal 973
Db 19438 CGCGCG-----CCGCCA 19427
Qy 974 SerProGlyValGlnProLeuTyrProIleProMetThrPro-----MetPro 989
Db 19426 CCG 19367
Qy 990 ValAsnGlnAlaLysThrTyrArgAlaValProAsnMetProGlnGlnArgGlnAspGln 1009
Db 19366 CCGCGCGCGCGGTTC-TCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19326
Qy 1010 HisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThr 1029
Db 19325 ---CACAGCGCGCGGTGGTCTCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19269
Qy 1030 ProProAlaTyr-SerThrGlnTyrVal----- 1038
Db 19268 CCGCGCTCACCAACACCGCTCGCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19209
Qy 1039 -----AlaTyrSerProGlnGlnPheProAsn 1047
Db 19208 ACCAGCGACGAGACCGCTCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19149
Qy 1048 GlnProLeuValGlnHisValProHisTyrGln---SerGlnHisProHisValTyrSer 1066
Db 19148 TCGCCCGCGCGCGCGTGTGATCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 19089
Qy 1067 ProValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGln----- 1083
Db 19088 GCGATGAGCTCGGCG 19029
Qy 1084 ---ProGlyLeuValSerSerSer----- 1090
Db 19028 CCGCGCGACGACGATGTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18969
Qy 1091 ---AlaThrGlnTyr-GlyAlaHisGluGlnThrHisAlaMetTyrAlaCysProLysLe 1109
Db 19096 CAGTGACCGGTACAGGT---CACGCTGCTCGCACCGCGCGCGCGCGCGCGCGCG 18912
Qy 1109 uProTyrAsnLysGluThrSerProSerPheTyrPheAlaIleSerThrGlySerLeuAl 1129
Db 18911 TCCG---CCGTCCCG 18855
Qy 1129 aGlnGlnTyrAlaHisProAsnAlaThrLeuHisProHisThrProHisProGlnProSe 1149
Db 18854 GCGTCGCGCGCGCGTGTGATCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 18795
Qy 1149 rAlaThrProThrGlyGlnGlnGlnSerGlnHisGlySerHisProAlaProSerPr 1169
Db 18794 ACG 18747
Qy 1169 oValGlnHisGlnHis-----GlnAlaAlaGlnAlaLeuHisLeuAlaSerPr 1186
Db 18746 GACCACCGAAGGCGATCGGGATGCCCTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 18687


```
QY 1186 oGlnGlnSerAlaIleTyrHis-----AlaGlyLeuAlaProThrPr 1201
Db 18686 CGCAGAGGGCTGGAGCGCCATCGAGCAAGCGCGGATGCAAGCCGAAGTCTTGC GCC 18627
QY 1201 oProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProAla----- 1219
Db 18626 CGC-TCCGAGACCCCTCGGGCAAGCCAGGTGCGGTAGATCGACCGCCCGCCACCGGCCA 18568
QY 1220 -AlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThr----- 1237
Db 18567 CGCCCGCGTCACGCCCGGCAACCGCGGCGGACCCACCGCTCGAGCGCTCGACCCACCG 18508
QY 1238 -----AsnProProHisMetAlaHisValProGlnAlaHis----- 1249
Db 18507 CQTGACGTGACCGCGCTCGGCCCGCGCGCGGCGGCGGCGCGCCCGCCCGCCACCGCGT 18448
QY 1250 -ValGlnSerGlyMetValProSerHisProThrAla-----His----- 1262
Db 18447 GCGCGCGGGCGGTGCGCCCGAGCAGCGCGCGCGCGGTGCGCGCGCGCGCGTCTCGT 18388
QY 1263 -AlaProMetMetLeuMetThrThrGlnProProGlyProGlnAlaAlaLeuAlaG1 1282
Db 18387 CGCGCGGTGCGCGCGGTGTGACCGCGAAGCCCGCGCGCGCGCGCGCGCGCGCGCTC 18328
QY 1282 nSerAlaLeuGlnProIleProValSerThrThrAla 1294
Db 18327 CACC-----GTCACCTGACGCGTCAACGCC 18303
```

RESULT 6

```
US-11-056-470-1
; Sequence 1, Application US/11056470
; Publication NO. US20050261263A1
; GENERAL INFORMATION:
; APPLICANT: SANTI, Daniel V.
; APPLICANT: MYLES, David C.
; APPLICANT: TIAN, Zong-Qiang
; APPLICANT: HUTCHINSON, C. Richard
; APPLICANT: JOHNSON, Robert G., Jr.
; APPLICANT: ZHOU, Yi-Qing
; APPLICANT: FENG, Li
; TITLE OF INVENTION: BENZOQUINONE ANSAMYCINS
; FILE REFERENCE: 30622007510
; CURRENT APPLICATION NUMBER: US/11/056,470
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 10/212,962
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/310,779
; PRIOR FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: US 60/389,255
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/393,929
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/395,275
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6210
; TYPE: DNA
; ORGANISM: Streptomyces geldanus
US-11-056-470-1
```

Alignment Scores:

Pred. No.:	3,74e-08	Length:	6210
Score:	435.00	Matches:	337
Percent Similarity:	32.98%	Conservative:	158
Best Local Similarity:	22.45%	Mismatches:	491
Query Match:	6.25%	Indels:	519
DB:	7	Gaps:	67

US-10-802-228-2 (1-1312) x US-11-056-470-1 (1-6210)

QY 9 ArgSerProAlaValAlaThrGluSerArgPheAlaAlaAlaArgTrpProGly--- 27

```
Db 101 CGATCGCGCGGT-----TCCGGAGGACCGGGGCTGGACCTGGCCCGGCTCT 148
QY 27 ----- 27
Db 149 TCGACTCGACCGGACCGCGTGGCAAGTCTATGTGCGCGAGGGCGGTTTCCTCACCG 208
QY 27 ----- 27
Db 209 ACGGGGCGGATTGACCGCGCATTTCTTCGGCATCTCGCCCGTGGCGGTGGCGATGG 268
QY 28 -----TTPArgSerLeuGlnArgProAla----- 35
Db 269 ACCCGCAGACGGTTGCTGTGAGACCGCGCTGGAGACCTTCCAGAATCCCGGAATCG 328
QY 35 ----- 35
Db 329 ACCCGGTTTCTCTGACCGGACCGACGTCGGTGTGTTTCAGCGGAGTGATGTACACGATT 388
QY 36 -----ArgArgSerGlyArgGly----- 41
Db 389 ACGGGGCGGACCGCGGACCGCGGAGGCGCTGGAGGGCATCTCGCGCTGGCGACGCG 448
QY 42 GlyGlyGlyAlaAlaProGlyProTyrProSer-----AlaAlaProProProGly 59
Db 449 CGGGAGCGTCTCTCGCGGCGGTGGCTACGGCGTGGGCTGACCGGCGCGCGTGA 508
QY 60 ProGlyProProSerArgGlnSerSerProProSerAlaSerAspCysPheGlySer 79
Db 509 CCGTGGACACCGCTGCTCGT-----CCTCCT--GGTAGCGTGCACCTGGCGGT 557
QY 80 AsnGlyAsnGlyGlyAlaPheArgProGlySerArgArgLeuLeuGlyGly 99
Db 558 TCAGCGGTGGCGACCGGCGGAATGCTCGTGGCGCTGCGCG-----GGGTGTCGCGGT 611
QY 100 ProProArgProPheValVal-----ValLeuLeuPro--LeuAlaSerProGlyAla 117
Db 612 GATGACGAGCGCGACGCTGCTTCATCGAGTTCTCCCGCGAGCGTGGCTCGCCCGCGG 671
QY 117 roProAlaAlaProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArgSerG 137
Db 672 CCGCTGCAAGTCTTCGCGAGGCGCGGACG---GCACCACTGGTCCGAGGGTGTGCG 728
QY 137 lyValSerLeuAlaArgProAlaProGlyCysProArgProAlaCysGluProValTyrG 157
Db 729 GTTGTGTGTC-----TGGAGCGGTCTCGATGCCCGCC----- 763
QY 157 lyProLeuThrMetSerLeuLys--ProGlnGlnGlnGlnGlnGlnGlnGlnGln 176
Db 764 -----GCAATGGGATGAGTGTCTCGCGCTGCTCGATGCCCGCC----- 815
QY 177 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 196
Db 816 CGGCGCCAGCAACCGCGCTGACCGCGCCCAACCGCGCGTCCCGAGAACGGGTGATCCGGCA 875
QY 197 LysProGly-GlySerGlyLeuLeuAlaSerProAlaAlaAlaProSerProSerSerSe 216
Db 876 GCGCTGCGGAACCGCGGCTGACGGTGGCGCGATGTGGACCGGTGAGAGGCCACCGGCAC 935
QY 216 rSerValSerSerSerAlaThrAlaProSerSerValValAlaAlaThrSerGlyG1 236
Db 936 GGGCAGCAGTCTCGGCGCACCGCATCGAGGCCCGAGGCGCTCCTG---GCCACTACCGGCA 992
QY 236 yGlyArgPro-----GlyLeuGlyArgGlyArgAsnSerAsnLysGlyLe 251
Db 993 GGAGCGCGCGGAGGATCAGCGCTGTGCTGCTGGG---TCGTTGAAGTCGAACATCGGCA 1049
QY 251 uProGlnSerThrIleSerPheAspGlyIleTyrAlaAsnMetArgMetValHisIleLe 271
Db 1050 TGCAGCGCGCGCGCGCGCGCGGTGTC-----ATCAAGATG----- 1089
QY 271 uThrSerValValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrGluG1 291
```

1090 Db -----GTCAGGCCATGCGCGCGGTACTGCCCAA 1121
291 QY yValPheLysThrTyrSerPro-----LysCysAspLeuValLeuAspAla----- 306
1122 Db AACCTCCACGCGACGAGCCACCAGCAAGGTGCACTGGAGTGGGTGCGT 1181
307 QY ----AlaHisGluLysSerThrGluSerSerGlyProLysArgGluGluMetG1 325
1182 Db ACTGTCCGAGCGCCCTGCGCGAGACGGGACACCCCGCGCGCGCGAATCTCTC 1241
325 QY uSerIleLeuPheLysCysSerAspPheValValGlnPheLysAspMetAspSerSe 345
1242 Db CTTGCGGTTCAGCGGACGACGACACAGCTGTCTGGAAACGGCACCCCTGGAGCGGC 1301
345 QY rTyrAlaLysArgAspAlaPheThrAspSerAlaIleSerAlaLysValAsnGluHi 365
1302 Db TGCACCCGAAACACAGCGGACGCGCGCGCTCTCGGCTCGTGGCCACGGGC----- 1356
365 QY sLysGluLysAspLeuProTrp----- 373
1357 Db -----GGGTAGTGGCGGTGTCTCCGCCAAGCTCTCGCGCGCTGGCGGC 1406
374 QY -AspAlaGlyGluLeuThrAlaAsnGluGluLeuGluAlaLeuGluAsnAspValSerAs 393
1407 Db TCAGGCAGAGCGTCTGTAGCCATCTGGAGTCCGGAGCGGACGCCAACCGGTGATGT 1466
393 QY nGlyTrpAspProAsnAspMetPheArgTyrAsnGluGluAsnTyrGlyValValSerTh 413
1467 Db GGGCTGG----- 1473
413 QY rTyrAspSerSerLeuSerSerTyrThrValProLeuGluArgAspAsnSerGluGluPh 433
1474 Db -----TCGCTGGCCACACCGCGCGCGGTGGAGCAC----- 1506
433 QY eLeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSerAlaGlnTy 453
1507 Db -----CGCGCGGTCTCTGGAGCGGATGCGAGGAGGATGGCGAC 1550
453 QY rLysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluGluLysTyrThrAlaVa 473
1551 Db GCGCGCGCTCTGGCG-----GAGGGCGCGCTGACCGCTCTGTGTACC----- 1596
473 QY lGlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGluAsnLys-TyrI 493
1597 Db -----GGACAGACCGG-AAACAGACGGCAAAACCGTGTTC 1630
493 QY leProProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGln-AsnSer 512
1631 Db TCTTCCCGCCCAAGCGCCCAATGGGTG-----GGCATGGGAGCCCAACTCTCTCA 1681
513 QY ---ProArgMetGlyGlnPro----- 518
1682 Db ACACCTACCGCTTTCGCGCGCGCTGCTGAGTGGCGGATGCTCTAGCGCGTATA 1741
519 QY -----GlySerGlySerMetProSerArg----- 526
1742 Db CCGACTGGTGGCTCATCGAGCTATCATCCGCGACCGCGCGCTCTCTCTGAGCGTG 1801
527 QY SerThrSerHisThrSerAspPheAsnProAsnSerGlySerAspGlnArgValValAsn 546
1802 Db TCGACGTCGTAC----- 1813
547 QY GlyGlyValProTrpProSerProCysProSerProSerSerArgProProSerArgTyr 566
1814 Db -----AGCCCGCCACTTGGCGGTCTGCTCTCC----- 1843
567 QY GlnSerGlyProAsnSerLeuProProArgAlaAlaThrProThrArgProProSerArg 586
1844 Db -----TCGCGGCACTCTGGCAATCCGTGGGCATCCACCCGACG---CGTCACTCGGCC 1894
587 QY ProProSerArgProSerArgProProSerHisProSerAlaHisGlySerProAlaPro 606
1895 Db ACTCCCAAGCGAAATCGCGCGCGCTGCGTGGCGGACACCTCACCTCACCAACGCGC 1954

607 QY ValSerThrMetProLys-ArgMetSerSerGluGlyProProArgMetSerProLys-- 625
1955 Db CCAAAATCGTCACTCCCTCGCGAGCCAGACATCGCCACACCTCGCGGACACGGCGCA 2014
626 QY ----AlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSe 644
2015 Db TGATGTCCCTCGCCACCCCGCGACACCATCG-ACCTCACCAACTGGCAGCGCAACTC 2073
644 QY rGlyLeu---GluPheValSerHisAsnProProSerGluAlaAlaThrPro---ProVa 662
2074 Db TGGATCGCGCACACAAACAGCCCAACCGCCATCGTCATCGAGCGACACCGACGCCCTG 2133
662 QY lAlaArgThrSerProSerGlyGlyThrTrpSerValValSerGlyValProArgLe 682
2134 Db CACCAACTCCACACCATACACCCAGGAGGACAGAGCCGCGCATCATCCCCGTGAC 2193
682 QY uSerProLysThrHisArgProArgSerProArgGlnAsnSerIleGlyAsn----- 699
2194 Db TAGGCTCCACACCGGACACGTTCGACACCATCAAAACACAGCTACAACAGCTACTCGAC 2253
700 QY ---ThrProSerGlyProValLeuAla-----SerProGlnAlaGlyI1 713
2254 Db GGCATCACCTCGAACCGGACCATCCCTCGCTCTCCACCGTCAACGAGCAGTGGATC 2313
713 QY eileProThrGluAla-----ValAlaMetProIleProAlaAlaSe 727
2314 Db GAACCAACACCGTGGGACAGCTACTGGTACCGCAACCTCCGCCAAACCGTGAATTC 2373
727 QY rProThrProAlaSerProAlaSerAsnArgAlaValThrProSerSerGluAlaLysAs 747
2374 Db GAGCACCATCTCCACACCTCGCCGACAGGGCTACCGCACCTACATCGAAATCA----- 2428
747 QY pSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnIleLysPr 767
2429 Db -----GCCCC 2433
767 QY oAsnGluThrSerProSerPheSerLysAlaGluAsnLysGlyIleSerProValValSe 787
2434 Db CACCCCGTCTCTCACACCGCCATCCAAAGAAACCTCGAAGCCCAACACACCCCAACA-- 2491
787 QY rGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsnAspPheArgLeuGlnPr 807
2492 Db -----CCACC 2496
807 QY oSerSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyLysSe 827
2497 Db ATCGTCACCGCA----- 2509
827 QY rArgAspLeuIleLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsnSe 847
2510 Db -----CCCTCCGCC----- 2518
847 QY rSerSerAsnCysThrSerGlySerSerLysProAsnSerProSerSerIleSerProSerI1 867
2519 Db -GGAGCAGACACCCCGCTCTCTCACCAACCTCGCCACCTCCACCAACA----- 2572
867 QY eLeuSerAsnThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSe 887
2573 Db -----ACGGAACAC-----CAGTCAAC 2589
887 QY rSerProAlaCysLysGlnGluLysAspAspLysGluGluLysAspAlaAlaGluG1 907
2590 Db TGGCCCAACCTCTTCACAGCA----- 2611
907 QY nValArgLysSerThrLeuAsnProAsnAlaLysGluPheAsnProArgSerPheSerGl 927
2612 Db -----CCCAACCCA-----CCGCATCCCCCTCCCC 2637
927 QY nProLysProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMe 947
2638 Db ACTTACCTCTCCACACACCACTACTGGCTCCCC---GCAACACAGCGGAGCGAT 2694

```
QY 947 t-----ValGlyHisGlnGlnProThrPro-----Va 956
Db 2695 GTGAGTGGCGTGGGCTCCAGGGCCAGCGGCCCTGGCGGGCGGTGTGAGCGTG 2754
QY 956 lTyrThrGlnProValCysPheAlaPro---AsnMetMetTyrProValProValSerPr 975
Db 2755 CCGGACACGGGGGTGTGTCTACCGCGGCGAGTGTGTGGTGCCCA-----2800
QY 975 oGlyValGlnProLeuTyrProIleProMetThrProMetProValAsn-----G1 992
Db 2801 -----CCACCCGTGGGTGGCGGACACCGCGCTCTCCGGAACGGTGTGTGTCGG 2850
QY 992 nAlaLysThrTyrArgAlaValProAsnMetProGlnGlnArg-----1006
Db 2851 GGCACCGCATGGCCGAATCCGCATCCGCGCGGAGAGACGACACCCCAACCCCTG 2910
QY 1007 -----GlnAspG1 1009
Db 2911 GAAGAGCTGGTCATCGGCCAGCGGATGACACTGCCCGAAGACGGTGCACTACATGTCAG 2970
QY 1009 nHisHisGlnSerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaTh 1029
Db 2971 GTACTGGTCGGCGGTGGAGGACGGCGCGGCGGAGGGTGGGATCTACTCTCGCCCGGAC 3030
QY 1029 rProPro-----AlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPh 1045
Db 3031 GCGGCCAGGAACAGGAATGGCTGGAGCACGCGCTCGG---GCACACTCGCCACGACGCG 3087
QY 1045 eProAsnGlnProLeuValGlnHisValProHisTyrGlnSerGlnHisProHisValTy 1065
Db 3088 GACGGTTCGGCGCGGGCGGCA-----TGGAGAACGGCATGCCCGAGTGGCGG 3135
QY 1065 rSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGlnProG1 1085
Db 3136 CCGCCCGGTGTCGAGCGCATGCTCTGGATGACTTCTACGACGACCTCGCCAGCGCGG 3195
QY 1085 yLeuValSerSerAlaThrGlnTyrGlyAlaHisGlnGlnThrHisAlaMetTyrAl 1105
Db 3196 TATG---AGTACGGCGCCGCTCCCGCGGCTGGAAGCGGTCTGGAAGCGCATGGCGAG 3252
QY 1105 aCys-----ProLysLeuProTyrAsnLysGluThrSerProSerPheThrPheAl 1122
Db 3253 GTGTTCGGGGGCGCGCTGCGGAGGAGACAGACGGAGCTCGCGGCC-----GGTTC 3306
QY 1122 alleSer-ThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrLeuHisProH 1142
Db 3307 GGTATCCATCCGGCGCTGTGGACGCGGTTGACGCGGCACTTCTGTGT-----3359
QY 1142 isThrProHisProGlnProSerAlaThrPro-----ThrGlyGlnGlnGlnSerGlnH 1160
Db 3360 --GCCCGCGGCGCGGCGCAACGCTCTCCCTTCGTGTGGAACGGGTACGGTGTGTG 3417
QY 1160 isGlyGlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaGlnA 1180
Db 3418 GCGGGCGGAGCACGCGCGTCCG-----TGTGCGCGCC 3450
QY 1180 laLeuHisLeuAlaSerProGlnGlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProT 1200
Db 3451 CCGGCCACCG-GCACGGACTCGTTCCAGCATCAGCC-----TGTTCGACAGCACCGG 3500
QY 1200 hrProProSer-----MetThrPro---AlaSerAsnThrGlnSerProGlnAsnSerP 1217
Db 3501 CTCCCGCGTGGCTCGGTGGACTCCCTGGTCTCCGGCGCATCAGTCCCGAGCAGCTCGC 3560
QY 1217 heProAlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrT 1237
Db 3561 TCCCG-----3565
QY 1237 hrAsnProProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProS 1257
Db 3566 -----CCTCGGCGGTGCGGTCGCTCGCTGATCGCTGTTACGCTGGAGTGAACCGA 3620
QY 1257 erHisProThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProG 1277
```

```
Db 3621 GCACCCCA-----CGCGCCTGGGACCGA 3644
QY 1277 lnAlaAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThr 1292
Db 3645 GGTTCCTGGGCCACCCCTCGGCGATG---CCACACCCAGCAGTGGACG 3688
RESULT 7
US-11-075-185-37/c
; Sequence 37, Application US/11075185
; Publication No. US2005026643A1
; GENERAL INFORMATION:
; APPLICANT: REEVES, CHRISTOPHER D
; APPLICANT: JULIEN, BRYAN
; APPLICANT: REID, RALPH
; TITLE OF INVENTION: BIOSYNTHETIC GENE CLUSTER FOR AMERUTICINS
; FILE REFERENCE: 010099, 03
; CURRENT APPLICATION NUMBER: US/11/075,185
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/551,103
; PRIOR FILING DATE: 2004-03-08
; PRIOR APPLICATION NUMBER: US 60/568,290
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 37
; LENGTH: 10524
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-11-075-185-37
Alignment Scores:
Pred. No.: 3,21e-07 Length: 10524
Score: 416.00 Matches: 324
Percent Similarity: 32.94% Conservative: 155
Best Local Similarity: 22.28% Mismatches: 602
Query Match: 5.98% Indels: 375
DB: 7 Gaps: 61
US-10-802-228-2 (1-1312) x US-11-075-185-37 (1-10524)
QY 5 AlaAlaAlaProArgSer-----ProAlaValAlaThrGluSerArgArgPhe 20
Db 5186 GCCCGCGGCCCGCGCAGCGGACGCGCGCCCTCCGCCCTGACAGAGTCGAGGACC 5127
QY 21 AlaAlaAlaArgTyrProGlyTyrArgSerLeuGlnArgProAlaArgArgSerGlyArg 40
Db 5126 GCCGGAAT-----CGTTCGTCTGGGGCCGTCGCGCCAGGAGCGCGCC 5082
QY 41 GlyGlyGlyAlaAlaProGlyProTyrProSerAlaAlaProProProGlyPro 60
Db 5081 AGCGCGAGGCTTGGCGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5022
QY 61 -----GlyProProProSerArgGlnSerSer 69
Db 5021 CGCACAGCTCGCAGCATCGCGAAGCGGCGCGCGCGCGCGCGCGCTCGCTCGAGC 4962
QY 70 ProProSerAlaSerAspCysPheGlySerAsnGlyAsnGly-GlyGlyAlaPheArgPr 89
Db 4961 CCGCGGAGATCGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTCC 4902
QY 89 oGlySerArgArgLeuLeuGlyLeu-----GlyGlyPr 100
Db 4901 AGCAGCTGAGCCCTTCTTGGCGGACATCGTCACGACACCCAGCGCTTCATCCGGCG 4842
QY 100 oProArgProPheValValValLeuLeuLeuProLeuAlaSerProGlyAlaProPro-----118
Db 4841 ATGTGCGCGC-----CGCGAGGTGCGCGCTCATGTCGCG 4809
QY 119 -AlaAlaProThrArgAlaSerProLeu-----GlyAlaArgAlaSerProProArgSe 136
Db 4808 AGCCAGGCTCGCGCGGAGAGCCCATGCGAGGCTCATCGCGGCGACCGCGCGCGCGCG 4749
```


Db 2904 GAAGACGACTGTCCGACAGCCAGCGGTGTCTCCCTGTGCAGCCGAC----- 2855
Qy 798 ysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuL 818
Db 2854 -----CGCCAGGATGTAGCTGTCCGTGCGCGGACCGCGT 2818
Qy 818 euAsnLysAsnArgGluGlyGluLysSerArgAspLeuLeuLysAspLysLeuGluPro 838
Db 2817 TCGCGCGCCAGCAGCGCGCGCGCGCTCG-----CGC 2785
Qy 838 erAlaLysAspSerPheLeuGluAsnSerSerAsnCysThrSerGlySerSerLysP 858
Db 2784 GAGCCGA-----GGTCGCTCGCTGATCCGCTCCTCGACG 2746
Qy 858 roAsnSer-----ProSerIleSerProSerIleLeuSerA 870
Db 2745 CTCGAGCCAGTAGCTGCGGTGGAACGCGTACGTGCGCAAGTCCACAGGTTGCCGCC 2686
Qy 870 snThrGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProA 890
Db 2685 GTACGGCGCGAGCACCCCTGGGCCACTCCACCGAATGCCCTGGACGTGACGCGCGCCAG 2626
Qy 890 laCysLysGlnGluLysAspAspLysGluLys-----LysA 903
Db 2625 CGCTGAALAAAGACCCCGCCCGTGTGCGCTGCAAGTGCACCGCCACACAGCC 2566
Qy 903 spAlaLaGluGlnValArgLysSerThrLeuAsnProAsnAlaLysGluPheAsnProA 923
Db 2565 GCCGCGCTCGGCGCTCGCGCGTCCAGCGGATGCCAGCACCGGATCGCGCTGACCTC 2506
Qy 923 rgSerPheSerGlnProLysProSerThrThrProThrSerProArgProGlnAlaGlnP 943
Db 2505 GACGAACACCCCGTCCCGCTCG-----CCAGACGCGGAGCGCGCGCTCGAG 2455
Qy 943 roSerProSerMetValGlyHisGlnGlnProThrProValThrThrGlnProValCysP 963
Db 2454 CGCACCGGCTCGCGAGTTCGCGCACAGTATGCC----- 2417
Qy 963 heAlaProAsnMetMetTyrProValProValSerProGlyValGlnProLeuTyrProI 983
Db 2416 -----CGCGAGCTCCGCGCGCGCGAGCACTTCTCCGT 2383
Qy 983 leProMetThr-----ProMetProValAsnGlnA 993
Db 2382 CACCGTCGATACAGCGGATGGCCAGCAGCGGGCTCGACCGAGCGCGAGCGGCTC 2323
Qy 993 laLysThrTyrArgAlaValProAsnMetProGlnGlnArgGlnAspGlnHisGlnS 1013
Db 2322 GAGCTCCGCGAGGATCGGTCCAGTGGCGCTGTGCAGCGGTAAATCGACGTGACCCG 2263
Qy 1013 erAlaMetMetHisProAlaSerAlaLaGlyPro-----ProIle-----A 1027
Db 2262 CCGCGGAAGACCTCTCTGTCTCCAGCTCGGCCAGCAAGCGATCGACCGTGTGCGCGCTC 2203
Qy 1027 laAlaThrProAlaTyrSerThrGlnTyrValAlaTyrSerProGlnGlnPhePro 1046
Db 2202 CCGGACACCAACCG-----TCGAGCCCGCGGTGTTCACCGCGCGCCAGCGACGTGCC 2149
Qy 1047 Asn-----GlnProLeuValGlnHisValProHisTyrGlnSerGlnHisProHis 1063
Db 2148 TCCGAGCGAGAGCCCTGCTC-----CACCTCGCGGACCGGGCGCTCGATCAGGCCAT 2092
Qy 1064 -ValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaProProThrHisAlaGl 1083
Db 2091 CGCGCGCGCGCGGATCGCGCGCAGCAGTGGCTGGCGAGCGCCAGACCGCGCGCGC 2032
Qy 1083 nProGlyLeuValSerSerAlaThrGlnTyrGlyAlaHisGluGlnThrHisAlaMe 1103
Db 2031 GTCCTCGAGCGTCAACGCCCGCGGACG----- 2004
Qy 1103 tTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerPheTyrPheAlaI 1123
Db 2003 -----TAGCGCGCGGACCTCGCCCTGGCTGTGCGCGCACG 1966

Qy 1123 eserThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAla-ThrLeuHisProHisT 1143
Db 1965 CGCGAAGGCTCGATG-----CCCCACGCCCGCCAGACCGCGGAG 1924
Qy 1143 hrProHisProGlnProSerAlaThrProThrGlyGlnGlnSerGlnHisGlyGlyS 1163
Db 1923 CCCACGCGCCATCGCAACAGCG---CCGGCTGCACCAACGTCACGCGGTCCGCGGGCGG 1867
Qy 1163 erHisProAlaProSer-ProValGlnHisGlnHisGlnAlaGlnAlaLeuHis 1182
Db 1866 CACGCGCTCTCTCTGCTCGCGCGGAGGACCGAGGACCGACCGCGCT----- 1816
Qy 1183 LeuAlaSerProGlnGlnSerAla-----IleTyrHisAlaGlyLeu 1197
Db 1815 ---GAACGCGCGACGCGCGGTGCGAGCGCGCGCGCGCGGCAAAACACTCGGACTC 1759
Qy 1198 AlaProThrProProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPhe 1217
Db 1758 GG-CCAGCAGCGCGCGCCCATCGCGGCACTGGCTCTCCCTGGCGCGGAG--- 1709
Qy 1218 ProAlaAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThr 1237
Db 1708 CCAGAACACGA----- 1697
Qy 1238 AsnProProHisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValProSer 1257
Db 1696 -----CCTTCGCGCGCGCGCGCTCGCGCGCGGTCCACCGGGGTGCGGCTCCCTT 1643
Qy 1258 HisProThrAlaHisAlaProMetMetLeuMetThrThrGlnProProGlyGlyProGln 1277
Db 1642 GCGAGAGCGCGCTCAGGCGCGCGCGCGCTCGCGCGCGCTCGCGCGCGTCAACCGCGCGC 1583
Qy 1278 -----AlaAlaLeuAlaGlnSerAlaLeuGlnProIleProValSerThrThr 1293
Db 1582 GCGCTCCAGGTGCTGCGCGCGCGCGCGCGCGGTGTGCACACCGCTCGCGCGCGCT 1523
Qy 1294 AlaHisPheProTyrMetThrHisProSer 1303
Db 1522 CCGGTGCGCGCGAGCT 1493

RESULT 8

US-11-121-086-4
; Sequence 4, Application US/1121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121.086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 164810
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-4

Alignment Scores:
Pred. No.: 7,54e-06 Length: 164810
Score: 407.00 Matches: 384
Percent Similarity: 31.84% Conservative: 156
Best Local Similarity: 22.84% Mismatches: 581
Query Match: 5.85% Indels: 583
DB: 78 Gaps: 78

US-10-802-228-2 (1-1312) x US-11-121-086-4 (1-164810)

Qy 2 ArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgPhe--- 20

QY 558 roSerSerArgPro-----ProSerArgTyrGlnSerGlyP 570
DB 87313 CCGGCCCCGGCCAGGAGGACGACTCTGCACTGACTCCGGGAGGACGCCGCTCAA 87372
QY 570 roAsnSerLeuProProArgAlaAlaThr----- 579
DB 87373 CTTCTCTAGCGCGCTCGCTCAGCGAGCAGACGTGAGGGACCCCGAGGACCGCC 87432
QY 580 -----ProThrArgProProSerArgProProSerArgP 594
DB 87433 CGGGGACACGCGGCGGCAAGACCCACCGCGCGCCCACT---CTGCAGACAGCG 87489
QY 594 roProSerHisProSerAlaHisGlySerProAlaProVal---SerThrMetProLysA 613
DB 87490 CATGGGACACCGGCACAAAGCGGGAGCGCGCGGCGAGCGGAGCAGTCTCGGGGCGC 87549
QY 613 rgMetSerSerGlu--- 620
DB 87550 GGGCAAGAACAGACGAGGCTGGAGCTGCCCTGGGCGCGCCCGCGAGCGCGCCGCGAGA 87609
QY 621 -----ArgMetSerProLysAlaG 627
DB 87610 CAAGGACGGCTCAAGCCCGCGGACCGCGGGGACCGGGGCTCCAGTCTGCTGCT 87669
QY 627 InArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSerGlyLeuG 647
DB 87670 CACGACGCGCCACTGAGGAGCGCTGTACTGCTTCTAC--GGCAACGACTCG----- 87718
QY 647 luPheValSerHisAsnProSerGluAlaAlaThrProProValAlaArgThrSer- 666
DB 87719 -----GACGAGAGCCCGCGCGCGCGGCCACGCCAACCCACCGCGGCGCATCGG 87770
QY 667 ProSerGlyGlyThrTrpSerSerValValSerGlyValProArgLeuSerProLysThr 686
DB 87771 CCATCC-----CTCGCGCTTTACGCGGAGCGTCCGCGAGCGCGGAGGAGG 87818
QY 687 HisArgProArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyPro----- 704
DB 87819 CCCCTGCGCCCT---CCAAGGCTGACCAAGTGTCCCGCGCGCGCGCGCGAGCCCA 87875
QY 705 -----ValLeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAlaMetPro 722
DB 87876 GCCTCATGTGACAGACCCCGCCCTGCT-----ACTCCCTGAGCTCCTCCGCA 87926
QY 723 IleProAlaAlaSerProThrProAlaSer-----ProAsnSerAsn 736
DB 87927 GCTCCCTCAGCGAGCCCGAGCCCTCGAGCCCGCGCGCGCTCCATCCACGAGCGCGGAGC 87986
QY 737 ArgAlaValThr-----ProSerSerGlu 744
DB 87987 CC-GGGGTACCAAGGACCCGGGCGCCAGGAGCGGACGACGAGTCTGCCCGCCGCGG 88045
QY 745 AlaLysAspSerArgLeuGlnAsp-----GlnArgGlnAsnSerPro 758
DB 88046 GCCCGGAGGAGCTTCTGCAGCGGTGATACAGTCCGCTCGCGCGCGCGCGCGCC 88105
QY 759 AlaGlyAsnLysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGlu 778
DB 88106 GTGTCTGCGCTCGCGCGCGCAAGCCCGAGCCACCGCGCTGGATGAGCGCGCGCGAG 88165
QY 779 -----AsnLysGlyIleSerProValValSerGluHisArgLysGlnIleAspAsp 795
DB 88166 GGGTCCCGGGAACGCGGCGGAGGAGCAGCGGCTCGGACCGGCGCTCCGAGCTGTATAGC 88225
QY 796 -----LeuLysLysPheLysAsnAspPhe----- 803
DB 88226 GTGGAGTGGCGGCGCATCCAGAGGGCGCCCAATTGTCACGTGCTGCACCCAGGCA 88285
QY 804 -----ArgLeuGlnProSerSerThrSerGluSerMetAspGlnLeuLeuAsn--- 819
DB 88286 GCAGTGCACCGCGGAGGCGCTCGTCCGAGTCCGACTCCATCTCTCTCTCTCTCTCTCT 88345

QY 820 -----LysAsnArgGluGluGlyLysSerArgAsp 829
DB 88346 CTGTCAGTGGGATCCACCTACAGCCCCCAAGCACAGGAAGGACGACGCGGAGGA 88405
QY 830 LeuIleLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsnSerSerSer 849
DB 88406 GAAATGGGCAGTGGCGGCGGCGGAGG-----GGCGCAGCCCTCA 88450
QY 850 AsnCysThrSerGlySerSerLys----- 857
DB 88451 GTCAGACACGCGGAGGCGCCCGGTTCCTCGACGCCGAGAGCCAGTGGCACACAG 88510
QY 858 -----ProAsnSerProSerIle-----Ser 864
DB 88511 AAGACACGCGCGGGTGCAGCTGTCTCCGGGACGAACAGTGTATCTACGTCCCCAGC 88570
QY 865 ProSerIleLeuSerAsnThrGluHisLysArgGlyProGluValThrSerGlnGlyVal 884
DB 88571 CCGCACCCCTGTGCCAGCCCAAGGAGACCCCGCGCGCCCGCCACACCGCGAAGGTG 88630
QY 885 -----GlnThrSerSerProAlaCysLysGlnGluLysAspAspLys 898
DB 88631 CGGCCCCCTTCTGCTGGCACAGCCCGCGCTCAGCC----- 88666
QY 899 GluGluLysLysAspAlaAlaGluGlnValArgLysSerThrLeuAsnProAsnAlaLys 918
DB 88667 -----AAAGTCCGAGCCCGCGGCGAGCGGTCTCGGAGGCTACACCGCGCTGCCAAG 88720
QY 919 GluPheAsnProArgSerPheSerGlnProLysProSerThrThrPro----- 934
DB 88721 ACCTCGGAGTGGGCGAGCTGAGCCACCCCGAAGAGCGCCACCGCCCGCGCTC 88780
QY 935 -----ThrSerProArgProGlnAlaGlnProSerProSerMetValGlyHisGlnGln 952
DB 88781 GCCAAGACCCCTCTCCAGCTCTCCAGACTCGCGCGC-----TCCAG 88828
QY 953 ProThrProValThrThrGlnProValCysPheAlaProAsnMetMetTyrProValPro 972
DB 88829 CCCCTGCCAGAAAGCGCCCGCTCACCCAGGCTGTGGGCGCTG----- 88876
QY 973 ValSerProGlyValGlnProLeuTyrProIleProMetThrProMetProValAsnGln 992
DB 88877 CCGGCGCCCGAGGCTCCCGGTCCCAAAAGCGCGCGCGCACCTCTCTCGCGAAGCAG 88936
QY 993 AlaLysThrTyrArgAlaValProAsnMetProGlnGlnArgGlnAspGlnHisGln 1012
DB 88937 CACAAGACGACAGATCGCCCGTGGATCCG----- 88964
QY 1013 SerAlaMetMetHisProAlaSerAlaAlaGlyProProIleAlaAlaThrProProAla 1032
DB 88970 ---TTCATGCAGAGCGCGCGCGGTGGCGCGCACCGCTGGGTGGGCGAGTCCCGAG 89026
QY 1033 TyrSerThrGln----- 1036
DB 89027 CCGGCGCCCGAGGCGCGGCGGAGCCGAGCGGCGCGCGCGCGCGCGCGCGCTG 89086
QY 1037 -----TyrValAlaTyrSerProGlnGlnPheProAsn-----GlnProLeuVal 1051
DB 89087 GGCCTGGTGGTGTGGCTC-AGCCCTCTCAGCGGCGAGCGAGTCTCCGACCGCTC- 89142
QY 1052 GlnHisValProHisTyrGlnSerGlnHisProHisValTyrSerProValIleGlnGly 1071
DB 89143 ---GGGCTTCCGGGACAGCTAACCTTCAT-----CAAGCA 89175
QY 1072 AsnAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuVal-SerSerSerAl 1091
DB 89176 GTCGCGGCGCTTGGCGCGCGCGCTCCGAGCTGTCTCGGCGGAGTCCGCGGCTCTGC 89235
QY 1091 aThrGlnTyrGlyAlaHisGluGlnThrHisAlaMetTyrAlaCysProLysLeuProTyr 1111
DB 89236 CCCCAG---GGCGCTCGCGCGCGCG-CCGCGCGCGCGCGCTGCCCGCGCTCTCTCTCT 89291
QY 1111 rAsnLysGluThrSerProSerPheThrPheAlaIleSerThrGlySerLeuAlaGlnGln 1131

Db 5376 CGT---GCTGCGCGCGTCTCTGTTACGCGCGAGCTGCGGATGACGCGCAG---CACGCG 5323
Qy 271 LeuThrSerValValGlySerLysCysGluValGlnValLysAsn----- 285
Db 5322 GTTCGCGGTTCGCGCGCGGTTCGAGCGCGTTCAGCGCCAGGATCCGCGACGCCCTCCGA 5263
Qy 286 -----GlyGlyIleTyrGluGlyValPheLysThrTyrSerProLysCys 300
Db 5262 CCAGCGCGCGCATCGCGCGCGCGCGAAGCGCTTGCAACGACGCTCAGCGCGGTGCG 5203
Qy 301 AspLeuVal-----LeuAspAlaAlaHisGluLysSerThrGluSerSer----- 315
Db 5202 GCGCTGGCGGTGAACATCGACGAAGATCATCGCGCTGGCCATCACGGTGGCGCGCGCGC 5143
Qy 316 -----SerGlyProLysArgGluGluIleMetGluSerIleLeuPhe 329
Db 5142 GAGGCCAGGTGCGACTCGCGCGCGCGGAGCATGCA----- 5104
Qy 330 LysCysSerAspPheValValGlnPheLysAspMetAspSerSerTyrAlaLysArg 349
Db 5103 -----CGCGAGGTGGACCGCGCACGACGACGA 5074
Qy 350 AspAlaPheThrAsp-----SerAlaIleSerAlaLysValAsnGlyGluHisLys 366
Db 5073 GCACGCGGTGTCGACGGTGACCGCGCGCGCGAGCCCCAGCGTGTAGCGCATCGCGCC 5014
Qy 367 GluLysAspLeuGluProTyrAspAlaGlyGluLeuThrAlaAsnGluGluLeuAla 386
Db 5013 CGAGCCGACGCTGAGGAA---GTTCCGCGTGGAGACGTGGCGCTCGAAGGACTCGAGCTG 4957
Qy 387 LeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMetPheArgTyrAsnGluGlu 406
Db 4956 CTG-----CAAGAG 4948
Qy 407 AsnTyrGlyValValSerThrTyrAspSerSerLeuSerSerTyrThrValProLeuGlu 426
Db 4947 CCGCCCGCGGTAGTCCCGTGAGCGAGCCGCGAAGACGCGCGTGGCGCTCGCCTCGAG 4888
Qy 427 ArgAspAsnSerGluGluPheLeuLysArgGluAlaArgAla---AsnGlnLeuAlaGlu 445
Db 4887 CGC-----GTGGGGCGAGGCGCGCGCGCTCGAGCGCTCCAGCGCCTCCAGCG 4846
Qy 446 GluIleGluSerSerAlaGlnTyrLysAlaArgValAlaLeuGluAsnAspAspArgSer 465
Db 4845 GCACCTCGAGCAGCAGCGCTGCTCGGGTCCATCGCTCCGCTTCGCGCGGCTCATGCC 4786
Qy 466 GluGluGluLysTyrThrAlaVal-----GlnArgAsnSerSer 478
Db 4785 GAAGAAAGTTCGATCGAAGCGATCGCGCTGCTAGGAGACGCGCGTGGTGATCGA 4726
Qy 479 GluArgGluGlyHisSerIleAsnThrArgGluAsnLysTyrIleProPro----- 495
Db 4725 CTTGCCCCGGGCGTCCGGGTGGGTGTGACGCCCGCGCACATCCAGCCCGCGCCCTC 4666
Qy 496 GlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgGlnAsnSerProArgMet 515
Db 4665 GGGGAACGGCC-CGA-----TCGCGTCCGCGCGCTCCGAGAGCA 4628
Qy 516 Gly-----GlnProGlySerGlySerMetProSerArg 526
Db 4627 GGTCCACACAGCCCGGGGGTATCGACGCCCGCGGGAGCGCGCACGCCATCGACACGA 4568
Qy 527 SerThrSer----- 529
Db 4567 TCGCATCGGTCTGTGTGCGCTCCCGCGCGCGCGGGCGCTGTGCGCGCTCGCCCCCG 4508
Qy 530 -----HisThrSerAspPheAsnProAsnSer----- 538
Db 4507 CGAGCTCCGAGAACGACGCTGAAGCAGCAGATCCGCGATCCGCGCGCGCTCGGATGT 4448
Qy 539 -----GlySer-----AspGlnArgValValAsnGlyGlyValProTrp 551
|||||
|||||

Db 4447 CGAAGCGAGGGTCTGTGGCAGCACCGTCTCGGGCGCGCGCGCGGCGGT----- 4397
Qy 552 ProSerProCysProSerProSerArgProProSerArgTyrGlnSerGlyProAsn 571
Db 4396 -----TCTTCAGCTCGAGCGCGGTGAGCGAATCGAGCCCCAGCTCTTCA 4352
Qy 572 SerLeuProProArgAlaAlaThrPro-----ThrArgProProSer 585
Db 4351 GCGGCTTTCGCGCGGACCTGCGCGGACCTGCGCGGACCTGCGCGGACGACGACCGCACCTCG 4292
Qy 586 ArgProProSerArgProSerArgProProSerHisProSerAlaHisGlySerProAla 605
Db 4291 CGCGCACGAGCTCGAGCAGCGCGCGCGCTCGGCTCGGGCA-----ACGCGGAGA 4238
Qy 606 ProValSerThrMetProLysArgMetSerSerGluGlyProProArgMetSerProLys 625
Db 4237 GCCGTGCGCGCAGCGCGCAGCGG-----CCGCTGCTCGGCTTCGG 4196
Qy 626 AlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIleSerSerGly 645
Db 4195 GCGTCTCGGCTCGCG-----TGGGCGCGCGCGGGC----- 4166
Qy 646 LeuGluPheValSerHisAsnProProSerGluAlaAlaThrProProValAlaArgThr 665
Db 4165 -----TGGCGCGCACCGCGCGCGCGCGCGGACCGCTCGGCGCTCGC 4118
Qy 666 SerProSerGlyGlyThrTrpSerSerValValSerGlyValProArgLeuSerProLys 684
Db 4117 TCGCGCGACGCTGGAGCG----- 4100
Qy 686 ThrHisArgProArgSerProArgGlnAsnSerIleGlyAsnThrProSerGlyProVal 705
Db 4099 -----CCGACAGGTCTGAGCGCGCGCGCACCA--- 4073
Qy 706 LeuAlaSerProGlnAlaGlyIleIleProThrGluAlaValAlaMetProIle---Pro 724
Db 4072 ---CGCTCGCTTCGCGGTGCGCGAGCGCGCTCGAGGAGCGCGGCGCTCGCGCGACCG 4016
Qy 725 AlaAlaSerProThrProAla-----SerProAlaSerAsnArgAlaValThrPro 741
Db 4015 ACAGGGGACAAAGGCCCTGCGCTCGAGCGGTGCGCGCGCGCGCGAGTTCGCGCG 3956
Qy 742 SerSerGluAlaLysAspSerArgLeuGlnAspGlnArgGlnAsnSerProAlaGlyAsn 761
Db 3955 TCATGCCAGCGCGCGCTCTCCCAAGACCCAGCGGAGCTGCTCGCGCGCACCGCGC 3896
Qy 762 LysGluAsnIleLysProAsnGluThrSerProSerPheSerLysAlaGluAsnLysGly 781
Db 3895 GCGCGCGAGCTGCGCGC---CGAACGCGTCCAGGAACGCAATTGGCAGCCCCGTAGTTTCG 3839
Qy 782 IleSerProValValSerGluHisArgLysGlnIleAspAspLeuLysLysPheLysAsn 801
Db 3838 CTTGCGCGCGGTTCGCCAGCGTACGAGATCGAG-GAGAACAGCACCAAGAGCCCGCAGC 3780
Qy 802 AspPheArgLeuGlnProSerThrSerGluSerMetAspGlnLeuLeuAsnLysAsn 821
Db 3779 GGCTGGTCTCTCGCGCGCGGTACACGTGGCAGCGCCCGCTCGACCTTC----- 3732
Qy 822 ArgGluGlyGluLysSerArgAspLeuIleLysAspLysIleGluProSerAlaLysAsp 841
Db 3731 ---GGCGCGCAGGGCGCG-----GTGAGCTGCTCGCGCTCGC 3696
Qy 842 SerPheIleGluAsnSerSerSerAsn-----CysThrSerGly----- 854
Db 3695 GCCGTGAGCACCCGTCCGCCAGCACCCCGCGGGTGTGCACACGCGCGCTCAGCGGGCGC 3636
Qy 855 -----SerLysProAsnSerProSerIleSerProSerIle 867
Db 3635 TTCGCGCGCGGACCTGGACTGCTGCCGAAAGCGCGCGCGCTCGTCACTCGTCCGACGCC 3576
Qy 868 LeuSerAsnThr-GluHisLysArg-GlyProGluValThrSerGln-----GlyValG 885
|||||
|||||
Db 3575 ACGACCTCGACCGTCTCGGCGCGCGTGGCGCGAGCTGTTCGACGAGCGCTCGGCGCTCC 3516


```
Db 137098 CGGGCGCGCCCGCTCCGCGCGCGACCTCGGGCGCGGAGACTCAGGCGGCGCGGG 137099
QY 87 Phe-----ArgProGlySerArg-----ArgLeuLeuGlyLeuGly 99
Db 137038 GCGGGCGCGCGGGGTCCGCCCGCGCGCAGACCCCGCGCGCGGACTGTGCGAGCGCGGG 136979
QY 100 ProProArgProPheValValValLeuLeuProLeuAlaSerProGlyAlaProProAla 119
Db 136978 CGCGCGCGCGCGAC-----CGCGCGGGGGCGCGCGCG 136943
QY 120 Ala-----ProThrArgAlaSerProLeuGlyAlaArgAlaSerProProArg--- 135
Db 136942 TCGGCGCTTTGTCGCGCGGAGGTGCGGA---GGGCGCGGGCGCCCTCCGCCCGTG 136886
QY 136-----SerGlyValSerLeuAlaArgProAlaProGlyCysPro 148
Db 136885 CCCCGCGCGCGGGGACTGGCGGGGTCTCCAGCTGGCGCGCGCCCTGCC----- 136832
QY 149 ArgProAlaCysGluProValTyrglyProLeuThrMetSerLeuLysProGlnGln 168
Db 136831 ---CGGGCC---CACCCGTCTCAGGC----- 136811
QY 169 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnPro 188
Db 136810 -----TCGTCCCGCGGGGCTGCCCGCAGCAGCCCGGATCTGCGCG 136769
QY 189 ProProAlaAlaAlaAsnValArgLysProGly-----GlySerGlyLeuLeu--- 204
Db 136768 CCCCCCGCGGCATCCGGGTGGCGCGCTCGGGTCTCCGAGGGCGGGCGCACCCCTG 136709
QY 205 -----AlaSerProAlaAlaAlaProSerProSerSerSer 216
Db 136708 CTTCCCGCGCGCCCGCGCTCCACCGCGCGGCTCGGGTCCGGGCTCTCACCTGC 136649
QY 217 SerValSerSerSerAlaThrAlaProSerValValAlaAlaThrSerGlyGly 236
Db 136648 CCGGTTCCAGCGCTATCCGAGCCTCACCAACCCCGCAGCAGCGCTCCAAATCGGGTC 136589
QY 237 GlyArg-----ProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGln 253
Db 136588 GGCACAACTCCAGTCTCCCTCGCGGGGCGGGGCGAGATTTTCAGAGGGCGCTGGCCAAAG 136529
QY 254 Ser-----ThrIleSerPheAspGlyIleTyrrAlaAsnMetArgMetVal 268
Db 136528 TCCAAAGCTGCTCGGGCTCGACCTCCCTGGGGCCCTCTCTGGGAGTCGCTCTCCGGACC 136469
QY 269 HisIleLeuThr-----Ser 273
Db 136468 CACGCTGTTCCTGGGGATGGAGTAAGTCCCTACACCTCAGCCCTTAAGGGTCCGGGAGC 136409
QY 274 ValValGly-SerLysCysGluValGlnValLysAsnGlyGlyIleTyrrGluGlyValPh 293
Db 136408 ATAGTAGGTGGACCATGTGGAGTGTGACAGCTGGGATGGATCCAAAGAGCC----- 136354
QY 293 eLysThrTyrrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrG1 313
Db 136353 -CAGAGTGAAGCCCG---GCCGACAGCAAAATCAGTCTCCAGCCAGCAATCTCGCAGGCA 136298
QY 313 uSerSerGlyProLysArgGluGluIleMetGlu-----SerIleLeuPhe-- 329
Db 136297 GGGGAGGCGCGTCCCGGGGCGACAGCTTGTGCAAAAGACCCGGGCTCTGATCTGTTTCA 136238
QY 329 ----- 329
Db 136237 GGAACCGCGAGTGATTCTCTCTCTGTTAGGTGACACAGGTGGGAACTGGGGCCCTGG 136178
QY 329 ----- 329
Db 136177 CCTCTGCTTTGTCTTTTTCCTCAATCTATAGACACAGTGTCTGGCCATGTTGGCCACGCTG 136118
QY 330 -----LysCysSerAspPh 334
Db 136117 GTCTTGAACCTCTCGGGCTCAGCGGATCTTCCCACTGGGCTTCCCAAAAGTGTAGATTA 136058
```

```
QY 582 gProProSerArgProProSerArg----- 590
Db 135049 GCCCCCT-----CCTGCTTCAAGGATCTCTGGGTGAGGAGGGGTGGCCAGGTCGG 134996
QY 591 -----ProSerArgProProSerHisProSerHisAlaHisGlySe 603
Db 134995 GAGGAGGGGTCTCTTCTCAACACGAGCTGGGACCTGCACAGGTGGGAGCTGGAAGGGC 134936
QY 603 rProAlaPro-ValSerThrMetProLysArgMetSerSerGluGlyProProArgMet- 622
Db 134935 ACCACCTCCACCCAGGCGATGGCAGGCGCACTATTTCATGTCAGCGGCGCCACCGAGCGAG 134876
QY 623 SerProLysAlaGlnArgHisProArgAsnHisArgValSerAlaGlyArgGlySerIle 642
Db 134875 AGC-----CGCCACCAACCCACCGCTGAAACGGATTGAGCAGAAGCTCCGT 134828
QY 643 SerSerGlyLeuGluPheValSerHisAsnProProSerHisAla----- 657
Db 134827 TTCCTTCTGTCTGTCTCTCTCAAGGTCAAGCCCAAGCAGGCGGACCAAGAGCGG 134768
QY 658 -----AlaThrProProValAlaAlaArgThrSerProSerGlyGlyThrTrpSer 673
Db 134767 TGGGTCCCCCACTCAACCTCCCGCGCGCTGCGCGCTCTAGAGGAGGACCTGGGA 134708
QY 674 SerValValSerGly-----ValProArgLeuSerProLysThrHisArgProArgSer 691
Db 134707 AGGAGGGG-GAGGCGCAGGGCGCTCTCGAGCGCTGCGCCTGGCGCAGGCTGCCCGGCAAA 134649
QY 692 ProArgGlnAsnSerIleGlyAsnThrProSerGlyProValLeuAlaSerProGlnAla 711
Db 134648 CCC-----CTGCTGGCAGGGGGGCCACTGTGGGGGTCCAGAA 134610
QY 712 GlyIleIleProThrGluAlaValAlaMetProIleProAlaAlaSerProThrPro-AL 731
Db 134609 GCCCTGTTCCTCTGCGGGCTGGAGCTGTCGGGCTTTGGAGCGCAGCGCGCGGAG 134550
QY 731 aSerProAlaSerAsnArgAlaValThrProSerSerGluAlaLysAspSerArgLeuG 751
Db 134549 GTCACTCTAGAGTAAGCTCTCTGTGGAGCA-----AGGGTCAG 134511
QY 751 nAspGlnArgGlnAsnSerProAlaGlyAsnLysGluAsnIleLysProAsnGluThrSe 771
Db 134510 AAATCAGGCTCACTCTCGG-----AACAGGTGTC 134481
QY 771 rProSerPheSerLysAlaGluAsnLysGlyIleSerProValValSerGluHisArgLy 791
Db 134480 GCCAAGTCAAGGGTTATCAGGGTCGCCCTCAGTCCTGTCTGGAGCCAGTACAGGC 134421
QY 791 sGlnIleAspAspLeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSe 811
Db 134420 T-----TGATGGGTCTGGCGCGCGCCCTGCTCT 134391
QY 811 rGluSerMetAspGlnLeuLeuAsnLysAsnArgGluGlyLysSerArgAspLeuI 831
Db 134390 GCACTGTGGAGCTGGGGTCTACCCAGCTACAGAGGGA----- 134351
QY 831 eLysAspLysIleGluProSerAlaLysAspSerPheIleGluAsnSerSerSerAsnCy 851
Db 134350 -----CTGTGCAGCCCGTG-----CCAGCAGCCCACTG 134322
QY 851 sThrSerGlySerSerLysProAsnSerProSerIleSerProSerIleLeuSerAsnTh 871
Db 134321 CCCACATGGC----- 134312
QY 871 rGluHisLysArgGlyProGluValThrSerGlnGlyValGlnThrSerSerProAlaCy 891
Db 134311 -----CAGGCGCAGGCTTGGGATCAGCTCCCGCAGCCCTCAG 134271
QY 891 sLysGlnGluLysAspAspLysGluLysLysAsp-----AlaAlaGluGlnValArgLy 910
Db 134270 CCTTCAAGGCGCGCAGCGAGGAGGAGCTTAAATAATGCGGGAGAGATTTCATTAC 134211
QY 910 sSerThrLeuAsnProAsnAla-----LysGluPheAsnProArgSerPheSerGlnProLy 929
```



```
QY 929 sProSerThrThrProThrSerProArgProGlnAlaGlnProSerProSerMetVal-- 948
Db 37576 CTCAGCCAGGCAAGCATCATGGCGGAGGGGTACAGGGGGCGTGGGGCGCGGAGGTGAC 37517
QY 949 -GlyHisGlnGln-----ProThrProValTyThrGlnPro----- 960
Db 37516 AGGCTCCACTAGGGGGAGGAGCCACCGCTGTCCATCGCCCCCAGCTGGGAAGCAGCG 37457
QY 961 -----ValCysPheAlaProAsnMetMetTyProValProValSerProGlyValG1 978
Db 37456 AGCTGGGTCTGGCTTAGCCCGAG-----GGGCCAGGGGGCG-- 37420
QY 978 nProLeuTyProLeuProMetThrProMetProValAsnGlnAlaLysThrTyArgAl 998
Db 37419 -----TTTCCCTGGGTGGGTGGCTCGCTCTCTGAAATAAAATGATCCCTTTCTCC 37367
QY 998 aValProAsnMetProGlnGlnArgGlnAspGlnHisGlnSerAlaMetMet-HisP 1018
Db 37366 AGTGGGTAAACATG-----AAAATCCCGAC 37343
QY 1018 roAlaSerAlaAlaGlyProProIleAlaAlaThrProAlaTySerThrGlnTyrv 1038
Db 37342 CCCCATCAGCATGAGCCAGATCTCTGCAACCCCTGCAGCCAGTGGTGAAGAAGAGC 37283
QY 1038 alAlaTySerProGlnGlnPheProAsnGlnProLeuValGlnHisValProHisTyrg 1058
Db 37282 AGGCACACAGCCCC-----ATTCCAGCTCACCC--TCCCTCCCTCCTCCCGCAGCTCCC 37232
QY 1058 InSerGlnHisProHisValTySer-----ProValIleGlnGlyA 1072
Db 37231 AFCCTTCCCATCTCCTCTTTCACACTCTCCTCCCTCCCTCCTCCCTCCTCCTCCT 37172
QY 1072 snAlaArgMetMetAlaProProThrHisAlaGlnProGlyLeuValSerSerAlaT 1092
Db 37171 CCTTCACACCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 37112
QY 1092 hrGlnTyR-GlyAlaHisGlnGlnThrHisAlaMetTyAlaCysProLysLeuProTy 1111
Db 37111 CCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 37055
QY 1112 AnLysGluThSer-----ProSerPheTyRphe 1121
Db 37054 TCCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 36995
QY 1122 AlAlleSerThrGlySerLeuAlaGlnGlnTyAlaHisProAsnAlaThrLeuHisPro 1141
Db 36994 GACCTTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 36938
QY 1142 HisThrProHisProGlnProSerAlaThrProThrGlyGlnGlnGlnSerGlnHisGly 1161
Db 36937 TTCACACCTTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 36899
QY 1162 GlySerHisProAlaProSerProValGlnHisGlnHisGlnAlaAlaGlnAlaLeu 1181
Db 36898 ---TCCACACCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 36842
QY 1182 -----HisLeuAlaSerProGlnGlnSerAla----- 1191
Db 36841 CTCGCCCTTCCCACTCTTCTCCCTCTTATATACCTCCTCCTCCTCCTCCTCCTCCTCCT 36782
QY 1192 IleTyHisAlaGlyLeuAlaProThrProPro-SerMetThrProAlaSerAsnThrG1 1211
Db 36781 CTTTATACCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 36722
QY 1211 nSerProGlnAsnSerPheProAlaAlaGlnGlnThrValPheThrIleHisProSerHi 1231
Db 36721 CCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 36689
QY 1231 sValGlnProAlaTyThrAsnProProHisMetAlaHisValProGlnAlaHisValG1 1251
Db 36688 C---CCTCCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 36640
```

```
QY 1251 nSerGlyMetValProSerHisProThrAla-HisAlaPro-MetMetLeuMetThrThr 1270
Db 36639 -----TCCCATCCTCTCTCTCCATTACATCTCTCTCTCTCTCTCTCTCTCTCTCT 36596
QY 1271 GlnProProGlyGlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGlnProIleProVal 1290
Db 36595 TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 36563
QY 1291 SerThrThrAlaHis-PhePro-----TyMetThrHisProSerValGlnAl 1306
Db 36562 TCTCCTCATCCACCTTTCAGTCTGGGGCCTGAACACAGACCCGCTTCTCCAGTGTCCAG 36503
QY 1306 aHisHisGlnGln 1310
Db 36502 ACACCACCGAGGT 36490
RESULT 12
US-10-821-234-218/c
; Sequence 218, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt SEQ_genes Version 1.0
; SEQ ID NO 218
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-218
Alignment Scores:
Pred. No.: 1,56e-06 Length: 7006
Score: 394.50 Matches: 379
Percent Similarity: 31.31% Conservative: 143
Best Local Similarity: 22.74% Mismatches: 599
Query Match: 5.67% Indels: 557
DB: 6 Gaps: 74
US-10-802-228-2 (1-1312) x US-10-821-234-218 (1-7006)
QY 1 MetArgSerAlaAlaAlaPro-----ArgSerProAlaValAlaThrGluSer 17
Db 6977 TTGGCAGCAGCAGCAGCAGCAACACACAGCAGCAGCAGCAGCGCGCGCGGAAACCCC 6918
QY 18 ArgArgPheAlaAlaAlaArgTrpProGlyTyRArgSerLeuGlnArgProAlaArgArg 37
Db 6917 GAGGCGCGCGCGCGCAGCCCGCGCTTCCCGA-----CGCGCAGCGC 6879
QY 38 SerGlyArgGlyGlyGlyValAlaAlaPro-----GlyProTyRPro-SerAlaAlaPr 55
Db 6878 GCGCGCAGCAGCAGCGGTGCGCGCGCGCGCTTAGCTCGCGCGCGCTCGCGCGCGGAC 6819
QY 55 oProProProGlyPro-----GlyProProSerArgGlnSerSerProProSe 72
Db 6818 CCCGACTCCCGGCCAGCAGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAGC 6759
QY 72 rAlaSerAspCysPheGlySerAsnGlyAsnGlyAlaPheArgProGlySerAr 92
Db 6758 GCGCGCAGAG-----CCCGGGGAGGGGGCGCGCA----- 6728
QY 92 gArgLeuLeuGlyLeuGlyGlyProArgProPheValVal----- 106
Db 6727 -----CCGCGCGCGCGCGCGATCTCTCAGGAGCCCGCAGAGCGCG 6690
```

QY 107 -----ValLeuLeuProLeuAlaSerProGlyAlaPro----- 117
Db 6689 GAGGGCGCGGCGCGCGAGCGGTGCTGGCCCGCGGGGCTCCCGGAGCTTCCCGCACCG 6630
QY 118 -----ProAlaAla-----ProThrAr 123
Db 6629 CTTGGGGCCGAGGAGCGGTGATCGGGCGGGCGCGCGCGCGCAAGGGTGGAGGAGCGG 6570
QY 123 gAlaSerProLeuGlyAlaArgAlaSerProProArgSerGlyValSerLeuAlaArgPr 143
Db 6569 CCG 6510
QY 143 oAlaProGlyCysProArgProAlaCysGluProValTyrglyProLeuThrMetSerLe 163
Db 6509 GCGGCTTAAGCG 6451
QY 163 uLySerProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 183
Db 6450 GCGGAGCG 6391
QY 183 n-----GlnGlnGlnGlnProProProAlaAlaAlaAlaAlaAlaAlaAlaAla 198
Db 6390 AGTGGCTACTCTGCGCAAGAGAGCATGGCCACAGCGCTTCTTCTGCTGCGCGGAC 6331
QY 198 roGlyGlySerGly-----LeuLeuAlaSerProAlaAlaAlaProSerPro----- 213
Db 6330 CCG 6271
QY 214 -----SerSerSerValSerSerSerAlaThrAlaProSerSerValV 230
Db 6270 TCAGTACTACGAGAGCGAGAAAGTGGCGGAGCAAGGCGCGCGCGCGCGCGCGCG 6211
QY 230 ala----- 231
Db 6210 TCGTCTCGACTGCTGCTGAACATCAACAAGCGCGCGCGCGCGCGCGCGCGCGCG 6151
QY 231 ----- 231
Db 6150 TCGCCTCTACCAAGGACGAGTACTTCGCGTGGCGCGCGCGAGAACGACGAGCAGG 6091
QY 232 -----AlaThrSerGlyGlyG 237
Db 6090 AGGGCTGTACCG 6031
QY 237 lyArgProGlyLeuGlyArgGlyArgAsnSerAsnLysGlyLeuProGlnSer----- 254
Db 6030 CGCCCCCG 5971
QY 255 --ThrIleSerPheAspGlyIleTyrrAlaAsnMetArgMetValHisIleLeuThrSerV 274
Db 5970 GCTGTGCG 5911
QY 274 alValGlySerLysCysGluValGlnValLysAsnGlyGlyIleTyrglyValPheL 294
Db 5910 CTTACCGTGGAGTGTGCGAGTGAACCTGAAGCCCAAGGCTCTG-----GGCCAGAGCA 5857
QY 294 yThrTyrrSerProLysCysAspLeuValLeuAspAlaAlaHisGluLysSerThrGlu 314
Db 5856 AGAACCTGACGGGGGTGACGCTGTGCGCTGTGCTGCG----- 5819
QY 314 erSerSerGlyProLysArgGluGluIleMetGluSerIleLeuPheLysCysSerAsp- 333
Db 5818 -----CGCACCATCGGCTTCGTGAAGCTCAACTCGCGAGCAGC 5782
QY 334 --PheValValGlnPheLysAspMetAspSerSerTyrrAlaLysArgAspAlaPhe 353
Db 5781 CGTCGGTGACCTGCAGCTCATGAACATC-----CGCGCTGCGCGGCACT 5737
QY 353 hrAspSerAlaIleSerAlaLysValAsnGlyGluHisLysGluLysAspLeuGluPro 373
Db 5736 CGGACAGCTTCTTCTTATCGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5677
QY 373 rp-----AspAlaGlyGluLeuThrAlaAsn-----GluGlu 384

Db 5676 GGATGACGCGGACGACTCGGTGGCGCAGACATCCACGAGACCATCTCGAGGCCA 5617
QY 384 euGluAlaLeuGluAsnAspValSerAsnGlyTrpAspProAsnAspMetPheArgTyra 404
Db 5616 TGAAGCGGCTCAAG-----GAGCTCTTCAGTTCC 5587
QY 404 snGluGluAsnTyrglyValValSerThrTyrrAspSerSerLeuSerSerTyrrValp 424
Db 5586 GGCCTGGC-----AGTAAGAGCAATCGTCGGGCTCGTCGGCCACGACC 5542
QY 424 roLeuGluArgAspAsn-----SerGluG 432
Db 5541 CCATCAGCTCCCG 5482
QY 432 luPheLeuLysArgGluAlaArgAlaAsnGlnLeuAlaGluGluIleGluSerSer----- 450
Db 5481 CGGGCTGTGGCGCGCTCGCGCACGACGCTGCGCGCGCGCGCGCGCGCGCGCGCG 5422
QY 451 --AlaGlnTyrrLysAlaArgValAlaLeuGluAsnAspAspArgSerGluGluLys 470
Db 5421 GCAGCTGTGCGCGGTGCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5362
QY 470 yrThrAlaValGlnArgAsnSerSerGluArgGluGlyHisSerIleAsnThrArgGlu 490
Db 5361 CGCGCGCGCGCGCGAGGCTGCTCGGTG-----GCTGGAGCGCGCGCGCGCGCG 5310
QY 490 snLysTyrrIleProProGlyGlnArgAsnArgGluValIleSerTrpGlySerGlyArgG 510
Db 5309 -----GCCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5264
QY 510 lnAsnSerProArgMetGlyGlnProGlySerSerMetProSerArgSerThrSerH 530
Db 5263 TCGCGCGCGCGCGCG-----GAGCAAGTGGCGCTGCTCGCGCGCGCGCGCGCG 5207
QY 530 isThrSerAspPheAsnProAsnSerGlySerAspGlnArgValValAsnGlyGlyValp 550
Db 5206 AG-----C 5204
QY 550 roTrpProSerProCysProSerProSerSerArgProProSerArgTyrglnSerGlyp 570
Db 5203 CGCT-CCATGTCTCATCGCGTGGCGACTCGCGCGCGCGCG-----CCACGACGCC 5154
QY 570 roAsnSerLeuProProArgAlaAlaThrProThrArgProProSerArgProProSerA 590
Db 5153 CGGCTCCCTGCTCTCCAGCAGCGCGCGCG-----GCTCGGGCTCTTACCG-CCGCGCGCG 5098
QY 590 rgProSerArgProProSerHisProSerAlaHisGlySerProAlaProValSerThrM 610
Db 5097 GCCCGCACCGCGCTCTGCGCGATCCGCTGCACACGCG-----CCGCGC----- 5053
QY 610 etProLysArgMetSerSerGluGlyProProArgMetSerProLysAlaGlnArgHisP 630
Db 5052 -----AGCGCGCTCCAGCGCGCGCGCGCTCCG-----C 5025
QY 630 roArgAsnHisArgValSerAlaGlyArgGlySerIleSerSerGlyLeuGluPheValS 650
Db 5024 CTGG----- 5020
QY 650 erHisAsnProProSerGluAlaAlaThrProProValAlaArgThrSerProSerGlyG 670
Db 5019 -----GCTCCCG 4965
QY 670 lyThrTrpSer-----SerValValSerGlyValProArgLeuSerProLysThrHisArgp 689
Db 4964 CGACCTCGCGCGCTCTGCGAGCCACCGAAGCAACACCGCGCGAGTCCATCCGCGAGACGCC 4905
QY 689 roArg-----SerProArgGlnAsnSerIleGlyAsnThrPro-----SerGlyp 704
Db 4904 CCGCGCGCGGAGACGCGCGCGCGCGCGGTGAGTTCTACGGGTATCATGATCATGACGAGGCC 4845
QY 704 ro-----ValLeuAlaSerProGlnAlaGlylle-----lleProT 716

Db 4844 CCTGAGCCACTGTGGCGCTCTCTACCGCGGGTCTCTGGGGGACGCGGCGCCAGGACCTGGA 4785
Qy 716 hrGluAlaValAlaMetProIleProAlaAlaSerProThrProAlaSerProAlaSerA 736
Db 4784 CCGAGGGCTGGCGAGA-----GGACCTACTCTCTGCTACCAAGCCAGCCAGCCGCGCAGCG 4734
Qy 736 snAtqAlaValThrProSerSerGluAlaLysAspSer---ArgLeuGlnAspGlnArg 755
Db 4733 GCCGGTGCCC-CAGCCCTCTCTGCTCGCTGGATGAATACACCCCTGATGCGGGCCACT 4675
Qy 755 lnAsnSerProAlaGlyAsnLysGluAsnLysProAsnGluThrSerProSerPheS 775
Db 4674 TCTCGGGCAGCGCGGCGCGCTCTG-CGCGTCTGCCCCCGGCTCTCTCCCAAGT----- 4620
Qy 775 erLysAlaGluAsnLysGlyIleSerProValSerGluHisArgLysGlnIleAspA 795
Db 4619 -----GGCCTACCACCCCTACCAGAGGACTACGGAGACATCGAGATC 4577
Qy 795 spLeuLysLysPheLysAsnAspPheArgLeuGlnProSerSerThrSerGluSerMetA 815
Db 4576 GGCTCCCAAG-----GAGCTCCAGCAGCAACCTCGGGGCGAGACGAGCGCTTACATG 4526
Qy 815 spGlnLeuAsnLysAsnArgGluGlyGluLysSerArgAspLeuIleLysAspLysI 835
Db 4525 CCCATGACCGCGCGCGCTTCCGGGCGAGTGGGAGCGGAGCTGCGAGGAGCGACAC 4466
Qy 835 leGluProSerAlaLysAspSerPheIleGluAsnSerSerAsnCysThrSerGlyS 855
Db 4465 TACATGCCCATGAGCCCGCCAGCGTGTCCGCC-CCAAGCAGATCTTGCAGCCCGAGGC 4407
Qy 855 erSerLysProAsnSer-----ProSerIleSerProSerIleLeuSerAenThrG 872
Db 4406 CGCGCCCGCCGCC-CCGCCCGCGTCTCTTGGCGGGCCTGCG----- 4364
Qy 872 luHisLysArgGlyProGluValThrSerGln----- 882
Db 4363 -----GGGCCAGCACCCACTCTTGGCGGGGCGAGACATTCGGGGGAGCGGGG 4315
Qy 883 --GlyValGlnThrSerSerProAlaCysLysGlnGluLysAspA 897
Db 4314 CGGGCTACAAGCCAGCTCGCCGCGGAGAGCTCCCGGAGGACAGTGGGTACATGCCCA 4255
Qy 898 -----LysGluGluLysLysAspAla----- 904
Db 4254 TGTGTGCGGTTCCAAGTGTTCATGGAGCATGCGAAGTGTGTGCCCAACGGGG 4195
Qy 905 -----AlaGluGlnValArgLysSerThr----- 912
Db 4194 ACTACCTCAACGTGTCCCGCGAGCGCGGTTCACACCGGGCACCCTCGCGGACTTCTTCT 4135
Qy 913 -----LeuAsnProAsnAlaLysLysPheAsn----- 921
Db 4134 CGCAGCCCTGACACCCCGCGGGAGCGCTCAGGGGCGTTCGGGCTGTCTACAGCT 4075
Qy 922 -----ProArgSerPheSerGlnPro-----LysProSerThr 933
Db 4074 CTTTGGCCCGCTCTTACAAGGCCCTCTACACCTGTGTGGGGGAGCAGCGACCACTACGTGC 4015
Qy 933 hrProThrSerPro-----ArgProGlnAlaGlnP 943
Db 4014 TCATGAGCTTCCCCGTGGGGCGCATCTCTGGAGGAGGCGTCTGGAGCCTCAGGCCACGC 3955
Qy 943 roSerProSerMetValGlyHis-----GlnGlnProThrProV 956
Db 3954 CAGGGCCCGCAGCGCGCGCGCTTCCGGGGCGGCCCCCAGCGAGCCCCCTCACCTG 3895
Qy 956 alThrThrGlnProVal---CysPheAlaProAsnMetMetTyrProValProValSerPr 975
Db 3894 TAGTGCCCTTCGCGCGGTAGCGGCGCGCGGTTCAGTTTCTTGGGGCCAGCGCGGCC 3835
Qy 975 oGlyValGlnProLeuProIlePro---MetThrProMetProValAsnGlnAlaLys 994
Db 3834 GGGCGGTGAGGCCACGCGCTGTCTCTCTGGAGGGGCTGCCCGACGCTGCCCGCATGCAACG 3775

Qy 994 sThrTyArgAlaValProAsnMetProGlnGlnArgGlnAspGlnHisHisGlnSerAl 1014
Db 3774 AGTACCACCTGCCACCGGAGCCCAAGAGCCCGGAGTACATCAACATCGACTT----- 3720
Qy 1014 aMetMetHisProAlaSerAlaAlaGlyProIleAlaAla-ThrProProAlaTyrS 1034
Db 3719 -----TGGCGAGCCCGGGCGCGCTGTGCGCCCGCGGCTCTCCCTGTGTGG 3673
Qy 1034 erThrGlnTyrValAlaTyrSer-ProGlnGlnPhePro-AsnGlnPro---LeuVal 1051
Db 3672 CGTCG-----GCGGCTCTCTCTCTCTCTCTCTGTCGCCCGCAGCGCCGCTGTGCGC 3619
Qy 1052 -----GlnHisValProHisTyrGlnSerGln----- 1060
Db 3618 TGGGCTCAGGCACCCCGGGCACACGACGACGACGCGGCGCGGTCTCCGCTCTCCGACT 3559
Qy 1061 ----HisProHisValTyrSerProValIleGlnGlyAsnAlaArgMetMetAlaPro P 1079
Db 3558 ACATGAACCTCGACTTCAGCTCCCCCAAGTCTCTTAAGCCGGGCGCCCCGAGCGGCCACC 3499
Qy 1079 roThrHisAlaGlnProGlyLeuValSerSerSerAlaThrGlnTyrGlyAlaHisGluG 1099
Db 3498 CCGTGGGCTCTTGGAGCGCTCTCTGTCCTCCCGGAGGCTCTCTC----- 3455
Qy 1099 lnThrHisAlaMetTyrAlaCysProLysLeuProTyrAsnLysGluThrSerProSerP 1119
Db 3454 -----CCGTATCCGCGTGTGCCCGCGCTCTCTTAAGCCGGGCGCCCCGCTCG- 3413
Qy 1119 heTyrPheAlaIleSerThrGlySerLeuAlaGlnGlnTyrAlaHisProAsnAlaThrL 1139
Db 3412 -----TCGTCTCT-GCAGCGCGCGCACCGCGCGCGCGCGCGCGCGGAGCTG 3368
Qy 1139 euHisPro-----HisThrProHisPro----- 1146
Db 3367 TACGCGCTGCCCGCGCTCGGCGTTCGCACCCCGCGGCGCGCGCGCGCTCATCG 3308
Qy 1147 ----- 1147
Db 3307 TTGTCTCGGACACCGGGACAATGGTGACTACACGAGATGGCTTTTGGTGTGCGCGCC 3248
Qy 1147 lnProSerAlaThr-----ProThrGlyGln----- 1155
Db 3247 ACCCGCGGCAACCTATCGCGGCCCCCGGAGCGCAGAGCTGCCCGCTGCCGACGCGCG 3188
Qy 1156 -----GlnGlnSerGlnHisGly----- 1162
Db 3187 AGTCGGGGGTGAAGAGGCTGAGCTCATGGAGCAGAGTGTGCGGAGTCGAGGCTTCTCTG 3128
Qy 1162 lySerHisProAlaProSerProValGlnHisHisGln---HisGlnAlaAlaGlnAlaL 1181
Db 3127 CAGGCCAGCCAGCCCGGACCCACCGCGCGCGCAAGGTTCATCCGCGCAGACCGCGCAG 3068
Qy 1181 euHisLeuAlaSerProGln---GlnGlnSerAlaIleTyrHisAlaGlyLeuAlaProT 1200
Db 3067 GGGGGCGCGCGCGCCACAGTTCCGAGACCTTCTCTCTCCACACGACGCGTCA-CCCCCGT 3009
Qy 1200 hrPro---ProSerMetThrProAlaSerAsnThrGlnSerProGlnAsnSerPheProA 1219
Db 3008 GTCCCCGTCTCTCCGCCACAAACCCCAAGCGGCCAACCTCGGCTCTCCGTGGGAAATGTCTC 2949
Qy 1219 laAlaGlnGlnThrValPheThrIleHisProSerHisValGlnProAlaTyrThrAsnPr 1239
Db 2948 TCTCAGGAAGACGAGCGCGCGGTGGTGTGCGGCTTGGAGGGGCGCAGCGCGCC 2889
Qy 1239 roPro-----HisMetAlaHisValProGlnAlaHisValGlnSerGlyMetValP 1256
Db 2888 CACCTCCCCACAGCATGTTCAGCGCGCGCGCTTGGCACCGGAGGCGCGCGCTGGAC 2829
Qy 1256 ro----- 1262
Db 2828 CCGGGTTCAGCCCGGGGCTTGGTGTCTGTCGGAGCGGTGGATCGCCCATCGCGCAG 2769

QY	1262	isAlaProMet	-----MetLeuMetThrThrGlnPro-----	1272
		:	:	
Db	2768	AGAGACCTCTCGCGGCTTCAGAAATGGTCTCAACTACATGCCCATCGACGTGAGGGAGGA	2709	
QY	1272	-----	-----	1272
Db	2708	GCCGGGCTGCCACCCAGCGGAGCGCGCGCGCGCTTCTCAGCGGGAGACAA	2649	
QY	1273	-----ProGly-----GlyProGlnAlaAlaLeuAlaGlnSerAlaLeuGln-----	1286	
		:	:	
Db	2648	GAGCTCTGGCGCGGAGCCGAAAGCCTCGGGGTCTCATCAGCGCTGTGGCGCTGCGGCAG	2589	
QY	1287	-----	-----ProIleProValSerThrT-----	1293
Db	2588	CACCGCGCGGGGTGCGGGGGCGGGTCCGGTCCCTGCCCTGCCAACACCTACGC	2529	
QY	1293	hrAla 1294		
Db	2528	CAGCA 2524		
RESULT 13				
US-11-087-100-1				
; Sequence 1, Application US/11087100				
; Publication No. US20050266440A1				
; GENERAL INFORMATION:				
; APPLICANT: Metz, James				
; APPLICANT: Barclay, William				
; APPLICANT: Flatt, James				
; APPLICANT: Kuner, Jerry				
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding ORFA of a PUFA Polyketide				
; FILE OF INVENTION: System and Uses Thereof				
; FILE REFERENCE: 2997-29				
; CURRENT APPLICATION NUMBER: US/11/087,100				
; PRIOR FILING DATE: 2005-03-21				
; PRIOR APPLICATION NUMBER: 09/231,899				
; PRIOR FILING DATE: 1999-01-14				
; PRIOR APPLICATION NUMBER: 60/284,066				
; PRIOR FILING DATE: 2001-04-16				
; PRIOR APPLICATION NUMBER: 60/298,796				
; PRIOR FILING DATE: 2001-06-15				
; PRIOR APPLICATION NUMBER: 60/323,269				
; PRIOR FILING DATE: 2001-09-18				
; NUMBER OF SEQ ID NOS: 37				
; SOFTWARE: Patent in version 3.3				
; SEQ ID NO 1				
; LENGTH: 8730				
; TYPE: DNA				
; ORGANISM: Schizochytrium sp.				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: (1)..(8730)				
US-11-087-100-1				
Alignment Scores:				
Pred. No.: 4.02e-06 Length: 8730				
Score: 386.00 Matches: 295				
Percent Similarity: 33.26% Conservative: 177				
Best Local Similarity: 20.79% Mismatches: 507				
Query Match: 5.55% Indels: 444				
DB: 7 Gaps: 53				
US-10-802-228-2 (1-1312) x US-11-087-100-1 (1-8730)				
QY	2	ArgSerAlaAlaAlaProArgSerProAlaValAlaThrGluSerArgArgPheAla 21	:	
Db	1223	CGCTCAAGCAACAAGACTCTCCGGGCAACATCAAGTCGACACCCACCACTCTTACG	1282	
QY	22	AlaAlaArg-----	-----TrpProGlyTrrp 28	
Db	1283	ACAACAGCGCCCATCAACGAGTCTCGCTCTACATTACACCATGAACCGCCCTGGTTCC	1342	
QY	29	ArgSerLeuGlnArgProAlaArgSerGlyGlyGlyGlyAlaAlaProGly 48		

[illegible]

This Page Blank (uspto)